CUK-H03251-77-0244805

77

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

WEI-LIANG CHOW
THE JOHNS HOPKINS UNIVERSITY

PHILIP HARTMAN
THE JOHNS HOPKINS UNIVERSITY

GILBERT A. HUNT
CORNELL UNIVERSITY

G. D. MOSTOW
YALE UNIVERSITY

WITH THE COÖPERATION OF

A. BOREL
S. S. CHERN
C. CHEVALLEY
K. L. CHUNG
J. A. DIEUDONNÉ
A. M. GLEASON
L. V. HÖRMANDER
K. KODAIRA
J. MILNOR
S. SMALE
A. WEIL

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LXXXV

1963

510.5

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

-IV

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

INDEX

ON L^p ESTIMATES AND REGULARITY, I.*[1]

By MARTIN SCHECHTER.

1. Introduction. In this paper we show how one may estimate the $H^{s,p}$ norms of solutions of elliptic boundary value problems for all *real* values of s and $p > 1$. These results are of particular interest when s is less than the order of the equation. From such estimates one obtains existence, regularity and representation theorems as well as other results. Details will be found in Sections 3 and 5.

Our estimates are of the form

$$\| u \|_{s,p} \leqq \text{const.} \left(\| Au \|_{s-m,p} + \| u \|_{s-m,p} \right)$$

for functions satisfying the homogeneous boundary conditions. Here A is an elliptic operator of order m; precise hypothesis are given in Sections 3 and 5. For s a non-negative integer, the norm $\| \ \|_{s,p}$ is merely the sum of the L^p norms of all derivatives up to order s. For s not an integer it is defined by means of Complex interpolation methods introduced by Calderón [9] and Lions [11] (cf. Section 2). For s negative, the norm is defined by duality.

As an example of a regularity theorem, we prove that if v is a distribution and

$$|(Au, v)| \leqq \text{const.} \| u \|_{m-s,p}$$

for all C^∞ function u satisfying the homogeneous boundary conditions, then v is in $H^{s,p'}$, where $p' = p/(p-1)$. We also obtain estimates for v. This generalizes a result of Agmon [2] for the Dirichlet problem.

From this, one easily obtains estimates of the form

$$\| u \|_{s,p} \leqq \text{const.} \left(\sum \| f_i \|_{s+m_i-m,p} + \| u \|_{s-m,p} \right)$$

for solutions of

$$Au = \sum A_i f_i,$$

where the A_i are differential operators of order $m_i \leqq m$. Also a very general representation theorem is obtained for bilinear forms which satisfy

$$a(u, v) = (Au, v)$$

for C^∞ function u, v satisfying homogeneous boundary conditions.

* Received December 18, 1961; revised July 5, 1962.

[1] Research for this paper was partly supported by the National Science Foundation Grant G-8205.

Some of these results were announced in the *Notices of the American Mathematical Society*, vol. 9 (1962), p. 40.

2. Interpolation spaces. We make use of the "complex" method of interpolation developed by Calderón [9, 10] and Lions [11]. Let X_0 and X_1 be Banach spaces and consider the set $H(X_0, X_1)$ of functions $f(x+iy)$ which are (a) continuous and bounded in the closed strip $0 \leqq x \leqq 1$ with values in $X_0 + X_1$, (b) analytic in $0 < x < 1$, and (c) such that $f(iy) \in X_0$, $f(1+iy) \in X_1$. Set

$$\| f \|_{H(X_0, X_1)} = \max[\sup_y \| f(iy) \|_{X_0}, \sup_y \| f(1+iy) \|_{X_1}].$$

For any θ such that $0 < \theta < 1$, the space $[X_0, X_1; \delta(\theta)]$ is the set of all $w \in X_0 + X_1$ such that $w = f(\theta)$ for some $f \in H(X_0, X_1)$. Provided with the norm

$$\| w \|_{[X_0, X_1; \delta(\theta)]} = \inf_{f(\theta)=w} \| f \|_{H(X_0, X_1)},$$

it becomes a Banach space.

An immediate consequence of the definition is

$$[X_1, X_0; \delta(\theta)] = [X_0, X_1; \delta(1-\theta)]. \tag{2.1}$$

We easily have in addition

LEMMA 2.1. (Calderón [9], Lions [11]). *If X_i and Y_i are Banach spaces, $i = 0, 1$, and T is a bounded linear mapping of X_0 into Y_0 and X_1 into Y_1, then it can be extended to be a bounded linear map of $[X_0, X_1; \delta(\theta)]$ into $[Y_0, Y_1; \delta(\theta)]$.*

The next two important results are due to Calderón [10].[2]

LEMMA 2.2. *If X_0 is a dense subspace of X_1, then it is dense in $[X_0, X_1; \delta(\theta)]$.*

LEMMA 2.3. *If X_0 and X_1 are reflexive, then*

$$[X_0, X_1; \delta(\theta)]' = [X_0', X_1'; \delta(\theta)].$$

Let G be a bounded domain in Euclidean n-space E^n with boundary ∂G of class C^∞.[3] Let $C^\infty(\bar{G})$ be the set of complex valued functions infinitely differentiable in the closure $\bar{G}$ of G. We set

---

[2] If X_0 and X_1 are reflexive, then Lemma 2.2 follows immediately from Lemma 2.3. For similar statements for other types of interpolation, cf. Lions, *Math. Scand.*, vol. 9 (1961), pp. 147-177.

[3] Some of our assumptions are made for convenience only and can be relaxed considerably.

$$\| u \|_{i,p} = [\sum_{|\mu| \leq i} \int_G | D^\mu u |^p \, dx]^{1/p} \tag{2.2}$$

for i a non-negative integer and $p > 1$. The summation is taken over all derivatives D^μ of order $|\mu| \leq i$. We let $H^{i,p}(G)$ denote the completion of $C^\infty(\bar{G})$ with respect to the norm (2.2).

For s positive and not an integer we define $H^{s,p}(G)$ to be

$$[H^{i,p}(G), H^{i+1,p}(G); \delta(\theta)], \qquad \theta = s - i,$$

where i is the greatest integer less than s. For s real and negative, $H^{s,p}(G)$ is defined as the completion of $C^\infty(\bar{G})$ with respect to the norm

$$\| u \|_{s,p} = \operatorname*{lub}_{v \in C^\infty(\bar{G})} \frac{|(u,v)|}{\| v \|_{-s,p'}}$$

where

$$(u,v) = \int_G u\bar{v} \, dx$$

and $p' = p/(p-1)$.

The following result is due to Lions-Magenes [12, III].

LEMMA 2.4. *If s_1 and s_2 are non-negative real numbers, then*

$$[H^{s_1,p}(G), H^{s_2,p}(G); \delta(\theta)] = H^{s_3,p}(G),$$

where $s_3 = (1-\theta)s_1 + \theta s_2$.

3. Boundary problems. Let A be a partial differential operator of order m with coefficients in $C^\infty(\bar{G})$.[4] If $C_0^\infty(G)$ denotes the set of those functions in $C^\infty(\bar{G})$ with compact support in G, let V be any space of functions such that

$$C_0^\infty(G) \subseteq V \subseteq C^\infty(\bar{G}).$$

We then take V' to be the set of those $v \in C^\infty(\bar{G})$ which satisfy

$$(u, A'v) = (Au, v)$$

for all $u \in V$, where A' is the formal adjoint of A. Clearly $C_0^\infty(G) \subseteq V'$. We define the following norms for s real and non-negative

$$| w |_{-s,p} = \operatorname*{lub}_{u \in V} \frac{|(w,u)|}{\| u \|_{-s,p'}}; \quad | w |'_{-s,p} = \operatorname*{lub}_{u \in V'} \frac{|(w,v)|}{\| v \|_{-s,p'}}$$

$$| w |_{s,p} = | w |'_{s,p} = \| w \|_{s,p},$$

where $p' = p/(p-1)$.

[4] Cf. footnote 3.

For $s \geqq 0$ we let $V^{s,p}(G)$ (resp. $V'^{s,p}(G)$) denote the closure of V (resp. V') in $H^{s,p}(G)$. For $s < 0$ we let these symbols denote the closures of $C^\infty(\bar{G})$ with respect to the norms $|\ |_{s,p}$ and $|\ |'_{s,p}$, respectively. Let N (resp. N') be the set of all $u \in V$ (resp. $v \in V'$) such that $Au = 0$ (resp. $A'v = 0$). For any set of functions S, the symbol S/N will denote those elements w of S which satisfy $(w, u) = 0$ for all $u \in N$. We let $H^{-\infty}(G)$ be the union of all the spaces $H^{s,p}(G)$, s real and $p > 1$.

We make the following assumptions for some fixed $p > 1$.

(a) There is a constant c_0 such that

$$\| u \|_{m,p} \leqq c_0(\| Au \|_{0,p} + \| u \|_{0,p}) \text{ for all } u \in V \tag{3.1}$$

$$\| v \|_{m,p'} \leqq c_0(\| A'v \|_{0,p'} + \| v \|_{0,p'}) \text{ for all } v \in V'. \tag{3.2}$$

(b) For every $f \in C^\infty(\bar{G})/N'$ there is a $u \in V$ such that $Au = f$, and for every $g \in C^\infty(\bar{G})/N$ there is a $v \in V'$ such that $A'v = g$.

It has been shown that assumptions (a) and (b) are satisfied for every $p > 1$ when A is a properly elliptic operator and V is determined by a set of $m/2$ differential operators of order $< m$ which cover A (cf. Section 5).

THEOREM 3.1. *For every real number s satisfying $0 \leqq s \leqq m$ there is a constant M_s such that*

$$\| u \|_{s,p} \leqq M_s(| Au |'_{s-m,p} + \| u \|_{s-m,p}) \text{ for all } u \in V \tag{3.3}$$

$$\| v \|_{s,p'} \leqq M_s(| A'v |_{s-m,p'} + \| v \|_{s-m,p'}) \text{ for all } v \in V'. \tag{3.4}$$

THEOREM 3.2. *If $u \in H^{-\infty}(G)$ and*

$$|(u, A'v)| \leqq c_1 \| v \|_{m-s,p'} \text{ for all } v \in V', \tag{3.5}$$

then $u \in V^{s,p}(G)$ and

$$\| u \|_{s,p} \leqq M_s(c_1 + \| u \|_{s-m,p}). \tag{3.6}$$

Similarly, if $v \in H^{-\infty}(G)$ and

$$|(v, Au)| \leqq c_1 \| u \|_{m-s,p} \text{ for all } u \in V, \tag{3.7}$$

then $v \in V'^{s,p}(G)$ and

$$\| v \|_{s,p'} \leqq M_s(c_1 + \| v \|_{s-m,p'}). \tag{3.8}$$

THEOREM 3.3. *If $a(u, v)$ is a bilinear form satisfying*

$$| a(u, v) | \leqq K_1 \| u \|_{s,p} \| v \|_{m-s,p'} \tag{3.9}$$

$$a(u, v) = (Au, v) \tag{3.10}$$

for all $u \in V/N$ and $v \in V'/N'$ then for every bounded linear functional F on $V^{s,p}(G)/N$ there is a $v \in V'^{m-s,p'}(G)/N'$ such that

$$Fu = a(u,v), \qquad u \in V^{s,p}(G)/N \tag{3.11}$$

For the Dirichlet problem Theorem 3.2 was proved by Agmon for s an integer. Theorem 3.1 was proved by Lions-Magenes [12, III] also for the Dirichlet problem when $N = N' = 0$, s real and satisfying $m/2 \leqq s \leqq m$. They employed Agmon's result.

Let A_i be a differential operator of order $m_i \leqq m$ with coefficient in $C^\infty(\bar{G})$, $1 \leqq i \leqq l$, where l is an integer $\geqq 1$.

THEOREM 3.4. *For every real number s such that $0 \leqq s \leqq m - \max m_i$, there is a constant C_s such that*

$$\| u \|_{s,p} \leqq C_s \Big(\sum_{i=1}^{l} \| f_i \|_{s+m_i-m,p} + \| u \|_{s-m,p} \Big) \tag{3.12}$$

for all u and f_i in $C^\infty(\bar{G})$ satisfying

$$(u, A'v) = \sum_{i=1}^{l} (f_i, A_i'v) \quad \text{for } v \in V'. \tag{3.13}$$

In particular, this holds if $u \in V$, $Au = \sum_{i=1}^{l} A_i f_i$, and

$$\sum_{i=1}^{l} [(f_i, A_i'v) - (A_i f_i, v)] = 0 \quad \text{for } v \in V'. \tag{3.14}$$

4. Proofs of the theorems. In proving the theorems of the previous section we make use of the important

THEOREM 4.1. *If s_1 and s_2 are non-negative real numbers, then*

$$[V^{s_1,p}(G), V^{s_2,p}(G); \delta(\theta)] \subseteq V^{s_3,p}(G) \tag{4.1}$$

with continuous injection, where

$$s_3 = (1-\theta)s_1 + \theta s_2.$$

Proof. Let X denote the space on the right hand side of (4.1). The identity mapping is bounded from $V^{s_1,p}(G)$ into $H^{s_1,p}(G)$ and from $V^{s_2,p}$ into $H^{s_2,p}(G)$. Hence, by Lemmas 2.1 and 2.4, it is a bounded mapping of X into $H^{s_3,p}(G)$. Thus $X \subseteq H^{s_3,p}(G)$ and

$$\| u \|_{s_3,p} \leqq K \| u \|_X \tag{4.2}$$

for all $u \in X$. Assume for definiteness that $s_1 < s_2$ (by (2.1) this entails no loss of generality). Since $V^{s_2,p}(G)$ is dense in $V^{s_1,p}(G)$, we see by Lemma

2.2 that it is dense in X. Thus for each element $u \in X$ there is a sequence $\{u_k\}$ of elements in $V^{s_2,p}(G)$ such that

$$\| u_k - u \|_X \to 0.$$

By (4.2) this sequence forms a Cauchy sequence in $H^{s_2,p}(G)$ and hence converges to an element $w \in V^{s_2,p}(G)$. Since the sequence approaches both u and w in $H^{s_1,p}(G)$, we have $u = w$ and the proof is complete.

Proof of Theorem 3.1. By (3.1)

$$\| u \|_{m,p} \leqq c_0 \| u \|_{0,p}$$

for all $u \in N$. Since G is bounded and ∂G is smooth, every closed bounded set in $H^{m,p}(G)$ is compact in $H^{0,p}(G) = L^p(G)$ (Rellich's lemma). Thus N is finite dimensional. This means that there is a constant c_3 such that

$$\| u \|_{m,p} \leqq c_3 \| Au \|_{0,p}$$

for $u \in V/N$. For otherwise there would exist a sequence $\{u_k\}$ of functions in V/N such that

$$(4.3) \qquad \| u_k \|_{m,p} = 1, \qquad \| Au_k \|_{0,p} \to 0.$$

By Rellich's lemma there is a subsequence which converges in $L^p(G)$. Denoting the subsequence again by $\{u_k\}$, we have by (3.1)

$$\begin{aligned} \| u_j - u_k \|_{m,p} &\leqq c_0 (\| A(u_j - u_k) \|_{0,p} + \| u_j - u_k \|_{0,p}) \\ &\to 0 \end{aligned}$$

as $j, k \to \infty$. Thus the u_k converge in $H^{m,p}(G)$ to an element $u \in V^{m,p}(G)/N$. By (4.3)

$$(4.4) \qquad \| u \|_{m,p} = 1$$

But by assumption (b), every function $w \in C^\infty(\bar{G})$ can be written in the form $w = A'v_1 + v_2$, where $v_1 \in V'$ and $v_2 \in N$. Hence

$$\begin{aligned} (u, w) &= (u, A'v_1) + (u, v_2) = \lim (u_k, A'v_1) \\ &= \lim (Au_k, v_1) = 0 \end{aligned}$$

for all $w \in C^\infty(\bar{G})$. Thus $u = 0$, contradicting (4.4). Hence (4.2) holds and similarly we have

$$(4.5) \qquad \| v \|_{m,p'} \leqq c_4 \| A'v \|_{0,p'}$$

for $v \in V'/N'$.

Next we claim that

$$(4.6) \qquad \| u \|_{0,p} \leqq c_5 \,| Au |'_{-m,p}$$

for all $u \in V/N$. This follows from the fact that

$$| Au |'_{-m,p} = \operatorname*{lub}_{v \in V'} \frac{|(Au, v)|}{\| v \|_{m,p'}} \geqq \operatorname*{lub}_{v \in V'/N'} \frac{|(u, A'v)|}{\| v \|_{m,p'}}.$$

Now for each $w \in C^\infty(\bar{G})/N$ there is a $v \in V'/N'$ such that $A'v = w$. Thus by (4.5)

$$c_4 \,| Au |'_{-m,p} \geqq \operatorname*{lub}_{w \in C^\infty(\bar{G})/N} \frac{|(u, w)|}{\| w \|_{0,p'}},$$

which is easily seen to be $\geqq$ const. $\| u \|_{0,p}$.

Thus by (4.2), (4.6) and assumption (b), we may consider the inverse operator A^{-1} as a continuous mapping of $L^p(G)/N'$ into $V^{m,p}(G)$ and $V'^{-m,p}(G)/N'$ into $L^p(G)$. Defining A^{-1} to be zero on N' we have, by Lemma 2.1, that A^{-1} is a bounded mapping of

$$X_\theta = [V'^{-m,p}(G), L^p(G)\,; \delta(\theta)]$$

into

$$Y_\theta = [L^p(G), V^{m,p}(G)\,; \delta(\theta)].$$

Thus

$$(4.7) \qquad \| u \|_{Y_\theta} \leqq c_6 \,\| Au \|_{X_\theta}$$

for all $u \in V/N$. Now by Theorem 4.1

$$(4.8) \qquad \| u \|_{s,p} \leqq c_7 \,\| u \|_{Y_\theta},$$

where $s = \theta m$. Moreover, since $V'^{m,p'}(G)$ is the dual space of $V'^{-m,p}(G)$ (cf. [17, Theorem 2.2]) we have by Lemma 2.3 that X_θ is the dual of

$$[V'^{m,p'}(G), L^{p'}(G)\,; \delta(\theta)] = [L^{p'}(G), V'^{m,p'}(G)\,; \delta(1-\theta)].$$

By Theorem 4.1, this last space is contained in $V^{m-s,p'}(G)$ with continuous injection. Thus its dual, X_θ, contains $V'^{s-m,p}(G)$ with continuous injection. Therefore we have

$$(4.9) \qquad \| w \|_{X_\theta} \leqq c_8 \,| w |'_{s-m,p}$$

for $w \in L^p(G)$. Combining (4.7), (4.8), and (4.9) we have

$$(4.10) \qquad \| u \|_{s,p} \leqq c_9 \,| Au |'_{s-m,p}$$

for $u \in V/N$. Now if $u \in V$, then $u = u' + u''$, where $u' \in V/N$ and $u'' \in N$. Thus

$$\| u \|_{s,p} \leqq \| u' \|_{s,p} + \| u'' \|_{s,p} \leqq c_9 \,| Au |'_{s-m,p} + c_{10} \,\| u'' \|_{s-m,p}$$

since N is finite dimensional (cf. [17, Lemma 7.1]). But

$$\| u'' \|_{s-m,p} \leqq \| u \|_{s-m,p} + \| u' \|_{s,p} \leqq \| u \|_{s-m,p} + c_9 \,|\, Au \,|'_{s-m,p}$$

and (3.3) follows immediately. The proof of (3.4) is similar and is omitted.

Proof of Theorem 3.2. By (3.5), $(A'v, u)$ is a bounded linear functional on $V'^{m-s,p'}(G)$. Thus there is an element $f \in V'^{s-m,p}(G)$ such that

$$(u, A'v) = (f, v) \tag{4.11}$$

and

$$| f |'_{s-m,p} \leqq c_1. \tag{4.12}$$

Clearly $(f, N') = 0$. Now there is a sequence $\{f_k\}$ of functions in $C^\infty(\bar{G})$ such that $| f_k - f |'_{s-m,p} \to 0$. Moreover for each k, $f_k = f_k' + f_k''$, where $(f_k', N') = 0$ and $f_k'' \in N'$. For each $w \in N'$

$$(f_k'', w) = (f_k - f_k', w) \to 0.$$

Since N' is finite dimensional, the f_k'' converge to zero in $L^2(G)$. Thus $| f_k' - f |'_{s-m,p} \to 0$. Now for each f_k' there is a $u_k \in V/N$ such that $Au_k = f_k'$. Thus by (4.10)

$$\| u_j - u_k \|_{s,p} \to 0$$

showing that they converge in $H^{s,p}(G)$ to an element $u_0 \in V^{s,p}(G)$. Moreover for $v \in V'$

$$(u_0, A'v) = \lim (u_k, A'v) = \lim (Au_k, v) = (f, v). \tag{4.13}$$

Subtracting (4.13) from (4.11) we have

$$(u - u_0, A'v) = 0.$$

But $u - u_0 = u' + u''$, where $(u', N) = 0$ and $u'' \in N$. Hence

$$(u', A'v) = 0$$

for all $v \in V'$. But if $w \in C^\infty(\bar{G})$, we know by assumption (b) that $w = A'v_1 + v_2$, where $v_1 \in V'$ and $v_2 \in N$. Hence

$$(u', w) = (u', A'v_1) + (u', v_2) = 0$$

showing that $u' = 0$. Thus $u = u_0 + u'' \in V^{s,p}(G)$ and the first assertion is proved. Inequality (3.6) follows from the fact that $u_k + u''$ converges to u in $V^{s,p}(G)$ while $A(u_k + u'') = f_k'$ converges to f in $V'^{s-m,p}(G)$. Since

$$\| u_k \|_{s,p} \leqq M_s(| Au_k |'_{s-m,p} + \| u_k \|_{s-m,p})$$

by (3.3), we have

$$\| u \|_{s,p} \leqq M_s(| f |'_{s-m,p} + \| u \|_{s-m,p}).$$

This, combined with (4.12), gives (3.6). The second part of the theorem is proved similarly.

Proof of Theorem 3.3. If F is a bounded linear functional on $V^{s,p}(G)/N$, then there is a $g \in V^{-s,p'}(G)/N$ such that

$$Fu = (u, g). \tag{4.14}$$

By the reasoning in the proof of Theorem 3.2, there is a $v \in V'^{m-s,p'}(G)/N'$ such that

$$(v, Au) = (g, u) \tag{4.15}$$

for all $u \in V^{s,p}(G)/N$. Since $(g, N) = 0$, this holds for all $u \in V^{s,p}(G)$. As in the proof of Theorem 3.2, one easily obtains a sequence v_k of functions in V/N which converges to v in $V'^{m-s,p'}(G)$. Thus

$$a(u, v) = \lim a(u, v_k) = \lim (Au, v_k) = (Au, v). \tag{4.16}$$

Combining (4.14), (4.15), and (4.16) we obtain the desired result.

Proof of Theorem 3.4. By (3.13),

$$\begin{aligned} | (u, A'v) | &\leqq \sum_{i=1}^{l} \| f_i \|_{s+m_i-m,p} \| A_i' v \|_{m-s-m_i,p'} \\ &\leqq c_{10} \| v \|_{m-s,p'} \sum_{i=1}^{l} \| f_i \|_{s+m_i-m,p} \end{aligned}$$

for all $v \in V'$, where we have made use of the fact that $m - s - m_i \geqq 0$ for each m_i (cf. [18, Lemma 4.3]). Setting

$$c_1 = c_{10} \sum_{i=1}^{l} \| f_i \|_{s+m_i-m,p},$$

we apply Theorem 3.2, which tells us that $u \in V^{s,p}(G)$ and that (3.6) holds.

5. Some applications. We now consider some situations when hypotheses (a) and (b) of Section 3 are known to hold. For such cases we can immediately conclude that Theorems 3.1-3.4 hold.

The operator A is *elliptic* in $\bar{G}$ if its characteristic polynomial $P(x, \xi)$ does not vanish for $x \in \bar{G}$, ξ real and not zero. It is *properly* elliptic in $\bar{G}$ if

for each boundary point x and each set of linearly independent vectors $\xi^{(1)}$ and $\xi^{(2)}$ the polynomial in z $P(x, \xi^{(1)} + z\xi^{(2)})$ has $m/2$ roots above and $m/2$ roots below the real axis (this assumes that m is even). Let $B_1, \cdots, B_{m/2}$ be differential operators of order less than m with coefficients in $C^\infty(\partial G)$. We say that they *cover* A if at each point $x \in \partial G$ the polynomials $Q_1(x, T + zN), \cdots, Q_{m/2}(x, T + zN)$ are linearly independent modulo the polynomial $(z - z_1) \cdots (z - z_{m/2})$ for each $T \neq 0$ tangent to ∂G at x and each $N \neq 0$ orthogonal to ∂G at x, where the $Q_j(x, \xi)$ are the characteristic polynomials of the B_j and the z_i are the roots of $P(x, T + zN) = 0$ with positive imaginary parts. Assume that A is properly elliptic in $\bar{G}$ and is covered by the B_j. If V is defined to be the set of these $u \in C^\infty(\bar{G})$ which satisfy

$$B_j u = 0 \text{ on } \partial G, \quad 1 \leqq j \leqq m/2, \tag{5.1}$$

then it has been shown that hypotheses (a) and (b) of Section 3 are true (cf. [4, 7, 8, 13, 14, 16, 19]). Thus we may conclude

THEOREM 5.1. *If A is properly elliptic in $\bar{G}$ and is covered by the B_j, then the conclusions of Theorems* 3.1-3.4 *are true.*

We next consider bilinear integro-differential forms of order r, $r \geqq 1$,

$$[u, v] = \int_G \sum_{|\mu|, |\tau| \leqq r} a_{\mu\tau}(x) D^\mu u \overline{D^\tau v}\, dx,$$

where the coefficients are in $C^\infty(\bar{G})$. Let $C_1, \cdots, C_r$ be a normal set of boundary differential operators of order $< r$. By this we mean that their orders are distinct and that ∂G is not characteristic to any of them at any point. Let U_h be the set of those $u \in C^\infty(\bar{G})$ which satisfy

$$C_j u = 0 \text{ on } \partial G, \quad 1 \leqq j \leqq h, \tag{5.2}$$

where h is a non-negative integer $\leqq r$. (If $h = 0$, $U_h = C^\infty(\bar{G})$). The bilinear form $[u, v]$ is said to be coercive on U_h if

$$\| u \|_{r,2}^2 \leqq \text{const.} (\operatorname{Re}[u, u] + \| u \|_{0,2}^2) \tag{5.3}$$

for all $u \in U_h$.

Since the set $\{C_j\}$ is normal of order $< r$, one has by integrating by parts

$$[u, v] = (Au, v) + \sum_{j=1}^r \int_{\partial G} F_j u \overline{C_j v}\, d\sigma \tag{5.4}$$

$$[u, v] = (u, A'v) + \sum_{j=1}^r \int_{\partial G} C_j u \overline{F_j' v}\, d\sigma, \tag{5.5}$$

where A and the F_j and F_j' are differential operators and A' is the formal adjoint of A (cf. [6]).

The following was proved by Agmon [1, 3].

LEMMA 5.1. *If $[u,v]$ is coercive over U_h then A is of order $2r$, is properly elliptic in G, and is covered by the operators $C_1,\cdots,C_h, F_{h+1},\cdots,F_r$.*

Moreover, one easily verifies in this case that the $\{F_j\}$ and $\{F_j'\}$ are normal sets of operators of orders $\geqq r$. If we let V be the set of those $u\in U_h$ satisfying

$$F_j u = 0 \text{ on } \partial G,\quad h<j\leqq r, \tag{5.6}$$

then V' becomes the set of those $u\in U_h$ which satisfy

$$F_j' u = 0 \text{ on } \partial G,\quad h<j\leqq r. \tag{5.7}$$

(cf. [6, 15]). Since the orders of the F_j and F_j' are $\geqq r$, we see that $V^{r,p}(G)=V'^{r,p}(G)=U_h^{r,p}(G)$ (cf. [18, Lemma 4.9]).

If $u\in N$, then $[u,v]=0$ for all $v\in U_h$ by (5.4). Conversely, if $u\in U_h$ and $[u,v]=0$ for all $v\in U_h$, then $u\in N$. Similarly, N' turns out to be the set of those $v\in U_h$ such that $[u,v]=0$ whenever $u\in U_h$. Thus we have

THEOREM 5.2. *Let $[u,v]$ be a bilinear form which is coercive over U_h. Let N be the set of those $u\in U_h$ which satisfy $[u,v]=0$ for all $v\in U_h$ and N' the set of those $v\in U_h$ which satisfy $[u,v]=0$ for all $u\in U_h$. Then every bounded linear functional Fu on $U_h^{r,p}(G)/N$ can be expressed in the form*

$$Fu=[u,v] \qquad u\in U_h \tag{5.8}$$

where $v\in U_h^{r,p'}(G)/N'$.

Necessary and sufficient conditions for $[u,v]$ to be coercive on U_h have been given by Agmon [1] (cf. also Schechter [14]). We consider here two simple cases.

COROLLARY 5.1. *Every bounded linear functional F on $H^{r,p}(G)$ can be represented by*

$$Fu=\sum_{|\mu|\leqq r}\int_G D^\mu u\,\overline{D^\mu v}\,dx, \tag{5.9}$$

where $v\in H^{r,p'}(G)$.

COROLLARY 5.2. *Let $H_0^{r,p}(G)$ denote the closure of $C_0^\infty(G)$ in $H^{r,p}(G)$. Then for every bounded linear functional F on $H_0^{r,p}(G)$ there is a $v \in H_0^{r,p'}(G)$ such that* (5.9) *holds.*

Corollaries 5.1 and 5.2 were known to several authors, including Milgram, Lions, Magenes, Szeptycki, Fritz and the author (cf. [12]).

THE UNIVERSITY OF CHICAGO
AND
NEW YORK UNIVERSITY.

REFERENCES.

[1] Shmuel Agmon, "The coerciveness problem for integro-differential forms," *Journal D'Analyse Mathematique*, vol. 6 (1958), pp. 183-223.

[2] ———, "The L_p approach to the Dirichlet problem, I," *Annali della Scuola Normale Superiore di Pisa*, Ser. 3, vol. 13 (1959), pp. 405-448.

[3] ———, "Remarks on self-adjoint and semi-bounded elliptic boundary value problems," *Proceedings of the International Symposium on Linear Spaces* held at the Hebrew University of Jerusalem, July 5-12, 1860, pp. 1-13.

[4] Shmuel Agmon, Avron Douglis, and Louis Nirenberg, "Estimates near the boundary of solutions of elliptic partial differential equations satisfying general boundary conditions, I," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 623-727.

[5] Nachman Aronszajn, "On coercive integro-differential quadratic forms," *Conference on Partial Differential Equations*, University of Kansas, 1954, Technical Report No. 14, pp. 94-106.

[6] Nachman Aroszajn and A. N. Milgram, "Differential operators on Riemann manifolds," *Rendiconti del Circolo Matematico di Palermo*, Ser. 2, vol. 2 (1953), pp. 1-61.

[7] F. E. Browder, "Estimates and existence theorems for elliptic boundary value problems," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 45 (1959), pp. 365-372.

[8] ———, "A priori estimates for solutions of elliptic boundary value problems, I, II, and III," *Koninklijke Nederlandse Akademie van Wetenschappen Indagationes Mathematicae ex Actis Quibus Titulus. Proceedings*, vol. 22 (1960), pp. 145-169.

[9] A. P. Calderon, Intermediate spaces and Interpolation, *Conference at Warsaw*, 1960.

[10] ———, "A paper on interpolation theorems," *to appear*.

[11] J.-L. Lions, "Une construction d'espaces d'interpolation," *Comptes Rendus des Séances de l'Academie Sciences*, Paris, vol. 250 (1960), pp. 1853-1855.

[12] J.-L. Lions and Enrico Magenes, "Problemi ai limiti non omogeni (III), (IV)," *Annali della Scuola Normale Superiore di Pisa*, Ser. 3, vol. 15 (1961), pp. 41-103, vol. 16 (1962), pp. 1-44.

[13] Jaak Peetre, "Another approach to elliptic boundary problems," *Communications on Pure and Applied Mathematics*, vol. 14 (1961), pp. 711-731.

[14] Martin Schechter, "Integral inequalities for partial differential operators and functions satisfying general boundary conditions," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 37-66.

[15] ———, "General value problems for elliptic partial differential equations," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 457-486.

[16] ———, "Various types of boundary conditions for elliptic equations," *Communications on Pure and Applied Mathematics*, vol. 13 (1960), pp. 407-425.

[17] ———, "On the theory of differential boundary problems," to appear (cf. *Bulletin of the American Mathematical Society*, vol. 66 (1960), pp. 495-500.)

[18] ———, "Coerciveness in L^p," *to appear.*

[19] L. N. Slobodeckii, "Estimates in L^p for solutions of elliptic systems (in Russian)," *Doklady Akademii Nauk SSSR*, vol. 123 (1958), pp. 616-619.

A CRITERION OF AN AMPLE SHEAF ON A PROJECTIVE SCHEME.*

By YOSHIKAZU NAKAI.[1]

In his previous paper [2],[2] the author proved a criterion of an ample divisor [3] on a non-singular surface, i. e., a divisor X on a non-singular surface F is ample if and only if $(X^2) > 0$ and X is arithmetically positive.[4] In this paper he will prove a generalization of this result to a projective scheme over a field. Originally the author intended to prove this generalization only for a non-singular variety of any dimension, say n. However, in the course of the proof it became necessary to treat the problem on a variety of dimension $n-1$ with singularities, and then on a scheme of dimension $n-2$. This is the reason why he finally decided to treat the problem on a general projective scheme right from the beginning.

The author wishes to express his heartfelt thanks to Professor O. Zariski and to D. Mumford. Without their encouragement and suggestions this work would not have been done.

In this paper we shall make extensive use of notations, terminologies, and the results in Grothendieck's book, "Éléments de Géométrie Algébrique," which will be cited as [G]. Most of them will be used freely without any further explanations, except some less fundamental notions. We hope the readers will not find much inconvenience in this way.

1. Morphism associated with invertible sheaf. Let S be a graded ring which is a homomorphic image of a polynomial ring over an algebraically closed field k. Let us denote by S_+ the set of elements of degrees > 0. Then the set of homogeneous prime ideals of S not containing S_+ with Zariski topology will be called a projective scheme over k and will often be denoted

* Received August 15, 1962.

[1] This work is supported by the Air Force Office of Scientific Research and Development, under Contract No. AF49(638)-494.

[2] The numbers in the brackets refer to the bibliography at the end of the paper.

[3] The meaning of the term "ample" is a little different from the classical usage (see the definition in Section 1). In his previous paper [2], the author used the term "non-degenerate divisor" to express this property.

[4] A divisor X on a non-singular surface F is called arithmetically positive if the Kroecker index $(X \cdot \Gamma)$ is always > 0 whenever Γ is a positive cycle. This nomenclature is due to O. Zariski. It was called (p)-divisor in [2].

by Proj (S). Let X be such a scheme, and let $\mathcal{L}$ be an invertible sheaf on X. Then $\Gamma(X, \mathcal{L})$ is a finite dimensional vector space over k (III-2.2.2).[5] Let us set $E = \Gamma(X, \mathcal{L}^{-1})$ and let $\mathcal{E}$ be the symmetric algebra of E over k (II-1.7.1). We shall denote as usual by $\mathcal{O}$, or $\mathcal{O}_X$, the structure sheaf of X. Then $\mathcal{E}$ defines in a natural way a sheaf of $\mathcal{O}$-module on X and we have an $\mathcal{O}$-homomorphism of graded algebras,

$$\psi : \mathcal{E} \to \mathcal{L}' = \sum_{n=0}^{\infty} \mathcal{L}^{\otimes -n}$$

From ψ we can define in a canonical way an open set $G(\psi)$ of X and a morphism

$$\gamma_{\mathcal{L}} : G(\psi) \to \mathrm{Proj}(\mathcal{E}) = Y$$

(II-§3). The morphism $\gamma_{\mathcal{L}}$ will be called the morphism associated with invertible sheaf $\mathcal{L}$.

We shall say that the sheaf $\mathcal{L}$ is *very ample* if $G(\psi) = X$ and $\gamma_{\mathcal{L}}$ gives an immersion of X onto Y and $\mathcal{L}$ is isomorphic to $\gamma_{\mathcal{L}}^*(\mathcal{O}_Y(1))$.

An invertible sheaf $\mathcal{L}$ will be called ample if $\mathcal{L}^{\otimes m}$ will become very ample for large m.[6] The morphism associated with an invertible sheaf $\mathcal{L}$ is regular at $x \in X$ if and only if there exists a section s in $\Gamma(X, \mathcal{L})$ such that $s(x) \neq 0$, i.e., $s(x) \notin \boldsymbol{m}_x \mathcal{L}_x$ where $\boldsymbol{m}_x$ is the maximal ideal of $\mathcal{O}_x$ (II-3.7.4).

THEOREM 1. *Retaining the notations and assumptions as above, let $\mathcal{L}$ be an invertible sheaf satisfying the following conditions:* (1) *$\gamma_{\mathcal{L}}$ is regular everywhere on X,* (2) *$\gamma_{\mathcal{L}}^{-1}(y)$ is a finite set for any $y \in Y$,* (3) *$\gamma_{\mathcal{L}}^*(\mathcal{O}_Y(1)) \cong \mathcal{L}$. Then $\mathcal{L}$ is an ample sheaf.*

Proof. Since any projective scheme is proper over k, $\gamma_{\mathcal{L}}$ is also a proper morphism (II-5.4.3). The condition (2) implies $\gamma_{\mathcal{L}}$ is a finite morphism by (II-4.2.2). On the other hand $\gamma_{\mathcal{L}}$ is dominate, hence $\gamma_{\mathcal{L}}$ is surjective, because any morphism which is at the same time dominate and proper must be surjective. The final assertion now follows from (III-2.6.2) since $\mathcal{O}_Y(1)$ is an ample sheaf on Y. Q.E.D.

2. The reduction of the problem.

Definition. Let $\mathcal{L}$ be an invertible sheaf on a projective scheme X. We

[5] This means the paragraph 2.2.2 of Chapter III of [G].

[6] This definition is different from the one given in [G], but they coincide as long as we treat projective scheme over a field (II-4.6.9).

shall say that $\mathcal{L}$ is arithmetically positive if for any reduced subscheme Y of positive dimension the Euler characteristic

$$\chi(\mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y) = \sum_{p=0}^{\infty} (-1)^p \dim H^p(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y)$$

increases indefinitely with m.

Our final object is to prove the

THEOREM 2. *Let X be a projective scheme over a field k and let $\mathcal{L}$ be an invertible sheaf on X. Then $\mathcal{L}$ is arithmetically positive if and only if $\mathcal{L}$ is ample.*

In this section we shall make some reduction of the problem. In the first place, we shall state the following proposition, which can be seen immediately from the definition.

PROPOSITION 1. *Let $\mathcal{L}$ be an arithmetically positive sheaf on X and let Y be any closed subscheme of X. Then the invertible sheaf $\mathcal{L}' = \mathcal{L} \otimes \mathcal{O}_Y$ is also an arithmetically positive sheaf on Y.*

PROPOSITION 2. *If Theorem 2 holds for reduced schemes, then it also holds for general schemes.*

Proof. Let X_{red} be the reduced scheme associated with X. Then we have the injection

$$j\colon X_{red} \to X.$$

Let $\mathcal{L}$ be an arithmetically positive sheaf on X. Then, since

$$j^*(\mathcal{L}) = \mathcal{L} \otimes \mathcal{O}_{X_{red}}$$

$j^*(\mathcal{L})$ is also an arithmetically positive sheaf on X_{red} by Proposition 1. The final assertion now follows from (II-4.5.14). Q.E.D.

PROPOSITION 3. *If Theorem 2 holds for any irreducible reduced scheme, then it also holds true for any reduced scheme.*

Proof. Let X_α $(\alpha = 1, 2, \cdots, s)$ be irreducible components of a reduced scheme X. Let π_α be the injection of X_α into X. Let Y be the disjoint sum $\coprod_\alpha X_\alpha$ of the schemes X_α (I-3.1), then we have a morphism

$$\pi\colon Y \to X$$

which is the sum of the morphisms π_α. As before we can see easily if $\mathcal{L}$ is arithmetically positive sheaf on X then $\pi_\alpha{}^*(\mathcal{L})$ is also an arithmetically

positive sheaf on X_α, *for* any α, hence $\pi^*(\mathscr{L})$ is also arithmetically positive on Y. By definition of Y, $\pi^*(\mathscr{L})$ is ample if and only if $\pi_\alpha^*(\mathscr{L})$ is ample on X_α for all α. Hence by assumption $\pi^*(\mathscr{L})$ is ample on Y. On the other hand, the morphism $\pi: Y \to X$ is surjective and finite as we can see easily from definition. The final assertion now follows from (III-2. 6. 2). Q. E. D.

Thus Theorem 2 is reduced to prove the

THEOREM 3. *Assume that Theorem 2 holds true for any* $(n-1)$-*dimensional projective scheme over* k. *Then it will also hold true for any* n-*dimensional irreducible, reduced scheme.*

We shall prove Theorem 3 in Section 3 $(n=1)$ and in Section 5 $(n>1)$.

In the following, an irreducible, reduced scheme over k will simply be called an irreducible variety over k.

3. Proof of Theorem 3 (the case $n=1$). Let X be a projective scheme over a field k and as before let $\mathcal{O}$ be the structure sheaf of X. Let $\mathcal{K}$ be the sheaf of total quotient ring of $\mathcal{O}$ on X. We shall denote by $\mathcal{O}^*$, $\mathcal{K}^*$, respectively, the sheaves of multiplicative groups composed of unit elements of $\mathcal{O}$ and $\mathcal{K}$ respectively. Then we have an exact sequence of sheaves of groups

$$(1) \to \mathcal{O}^* \to \mathcal{K}^* \to \mathcal{K}^*/\mathcal{O}^* \to (1).$$

From this we have an exact sequence of cohomology groups

$$k \to H^0(X, \mathcal{K}^*) \xrightarrow{\psi} H^0(X, \mathcal{K}^*/\mathcal{O}^*) \xrightarrow{\delta} H^1(X, \mathcal{O}^*)$$
$$\to H^1(X, \mathcal{K}^*) \to$$

As is well known, an invertible sheaf is represented, up to isomorphisms, by an element of cohomology group $H^1(X, \mathcal{O}^*)$. An element of $\boldsymbol{D} = H^0(X, \mathcal{K}^*/\mathcal{O}^*)$ defines in a canonical way an invertible sheaf, called a divisor, on X. The divisor defined by $\boldsymbol{P} = \mathrm{Im}(\psi)$ form a subgroup of $\boldsymbol{D}$, called the group of principal divisors. If X is an irreducible variety over k, the sheaf K^* is a constant sheaf and hence $H^1(X, \mathcal{K}^*) = 0$. Then the homomorphism δ is surjective. It implies that any invertible sheaf is represented by a divisor on X.

Remark. More generally, $H^1(X, \mathcal{K}^*) = 0$ if X is a reduced projective scheme over k. If X is not a reduced scheme, it cannot be expected that we have $H^1(X, \mathcal{K}^*) = 0$. Nevertheless, the homomorphism δ is surjective if X

is a projective scheme defined over an infinite field.[7] (Cf. forthcoming paper [3].)

A divisor $\mathcal{D}$ is, by definition, a subsheaf of $\mathcal{K}$. We shall say that $\mathcal{D}$ is positive (or integral) at a point $x \in X$ if $\mathcal{D}_x \subseteq \mathcal{O}_x$. A divisor $\mathcal{D}$ is called a positive (or an integral) divisor if $\mathcal{D}$ is positive everywhere on X. The set of points x on X such that $\mathcal{D}_x \neq \mathcal{O}_x$ will be called the support of $\mathcal{D}$, and will be denoted by $\mathrm{Supp}\{\mathcal{D}\}$. It is easily seen that a divisor $\mathcal{D}$ can be represented as a quotient of two integral divisors, though there is no canonical expression for it unless X is everywhere normal.

Now assume that X is one dimensional scheme over k and let $\mathcal{A}$ be an integral divisor on X. Then we can define the notion of degree, $\deg \mathcal{A}$ in symbol, by

$$\dim_k H^0(X, \mathcal{O}/\mathcal{A}) = \sum_{x \in \mathrm{Supp}(\mathcal{A})} \mathrm{length}\,(\mathcal{O}_x/\mathcal{A}_x).$$

Let $\mathcal{D}$ be an arbitrary divisor and let $\mathcal{D} = \mathcal{A} \otimes \mathcal{B}^{-1}$ be a representation of $\mathcal{D}$ as a quotient of two integral divisors. Then the degree of $\mathcal{D}$ is defined by

$$\deg \mathcal{D} = \deg \mathcal{A} - \deg \mathcal{B}.$$

If f and g are two non-zero divisors of $\mathcal{O}$ we have

$$\mathrm{length}(\mathcal{O}/(fg)) = \mathrm{length}(\mathcal{O}/(f)) + \mathrm{length}(\mathcal{O}/(g)).$$

From this we see easily that the definition of $\deg \mathcal{D}$ does not depend on any particular representation as a quotient of integral divisors.

Let $\mathcal{L}$ be a divisor on a projective scheme over k. We define $\chi(\mathcal{L}^{-1})$ by

$$\chi(\mathcal{L}^{-1}) = \sum_{p=0}^{\infty} (-1)^p \dim_k H^p(X, \mathcal{L}^{-1}).$$

This is a well-defined integer, since these cohomology groups are of finite dimensional over k (III-2.2.2).

PROPOSITION 4 *Let X be a one-dimensional projective scheme over a field k and let $\mathcal{L}$ be a divisor on X. Then we have*

$$\chi(\mathcal{L}^{-1}) = \deg(\mathcal{L}) + \chi(\mathcal{O})$$

The proof is quite similar to the proof given in [4], where X is an irreducible curve, and it will be omitted.

COROLLARY. *The degree of principal divisor is* 0.

Proof. Let t be a section of $H^0(X, \mathcal{K}^*)$. Then the sheaf $\mathcal{O}t$ is isomorphic to $\mathcal{O}$ and hence $\chi(\mathcal{O}) = \chi(\mathcal{O}t)$.

---

[7] This will be proved in a more general situation in [3].

In the rest of this paragraph, we shall assume X is an irreducible variety over k.

PROPOSITION 5. *Let $\mathfrak{D}$ be a divisor of positive degree on X, then $H^1(X, \mathfrak{D}^{\otimes -m}) = \{0\}$ if m is big enough.*

Proof. There exists a divisor $\mathfrak{D}_1$ linearly equivalent to $\mathfrak{D}$ such that any point in Supp $\{\mathfrak{D}_1\}$ is simple on X. Since $H^1(X, \mathfrak{D}^{\otimes -m}) = H^1(X, \mathfrak{D}_1{}^{\otimes -m})$, it suffices to prove under the assumption that any point in Supp $\{\mathfrak{D}\}$ is simple on X. This is proved in [4].

PROPOSITION 6. *Let $\mathfrak{D}$ be a divisor of positive degree on X. Then for large m the morphism γ_m associated with the divisor $\mathfrak{D}^{\otimes m}$ is everywhere regular on X.*

Proof. Let x be an arbitrary point of X algebraic over k and we shall show that there exists an integer m such that γ_m is regular at x. Since the domain of regularity is an open subset of X it suffices for the proof of the proposition. Without loss of generality we can assume that x is not in the support of $\mathfrak{D}$, i.e., $\mathfrak{D}_x = \mathcal{O}_x$. Let f $(\neq 0)$ be an arbitrary non-unit element of $\mathcal{O}_x$ and let $\mathfrak{f}$ be the divisor defined by $\mathfrak{f}_x = \mathcal{O}_x f$ and $\mathfrak{f}_y = \mathcal{O}_y$ if $y \neq x$. We shall consider an exact sequence of sheaves

$$0 \to \mathfrak{D}^{\otimes -m} \otimes \mathfrak{f} \to \mathfrak{D}^{\otimes -m} \to \mathfrak{D}^{\otimes -m} \otimes (\mathcal{O}/\mathfrak{f}) \to 0.$$

Since $(\mathcal{O}/\mathfrak{f})_y = 0$ if $y \neq x$ and $\mathfrak{D}_x = \mathcal{O}_x$, $\mathfrak{D}^{\otimes -m} \otimes (\mathcal{O}/\mathfrak{f}) = \mathcal{O}_x/(f)$. On account of Proposition 5 we have the exact sequence

$$H^0(X, \mathfrak{D}^{\otimes -m}) \to \mathcal{O}_x/(f) \to 0.$$

Let s be a section of $H^0(X, \mathcal{L}^{\otimes -m})$ such that $s_x \equiv 1 \bmod (f)$. Then s_x is a unit of $\mathcal{O}_x$ and γ_m is regular at x, completing the proof. Q. E. D.

Proof of Theorem 3. In the case where $\dim X = 1$, the condition (2) of Theorem 1 is trivial. As to condition (3), we proceed as follows. Without loss of generality, we can assume that $\mathfrak{D}$ is a positive divisor. Then 1 is a section of $H^0(X, \mathfrak{D}^{\otimes -m})$. By condition (1) for any point x on X, there exists a section s in $H^0(X, \mathfrak{D}^{\otimes -m})$ such that $s(x) \neq 0$. Hence we have $\mathfrak{D}_x{}^{\otimes -m} = \mathcal{O}_x s$. Since s is in the affine coordinate ring of $Y = \gamma_m(X)$, it implies the condition (3).

Thus all conditions in Theorem 1 are satisfied for a divisor $\mathfrak{D}^{\otimes -m}$ for large m, hence Theorem 3 is proved in the case where $n = 1$. Q. E. D.

4. Fundamental lemmas on ample sheaves.

PROPOSITION 7. *Let $\mathcal{L}$ be an ample sheaf on a projective scheme X over a field k. Then for any coherent sheaf $\mathcal{F}$ on X we have $H^{\rho}(X, \mathcal{F} \otimes \mathcal{L}^{\otimes -m}) = \{0\}$, for a sufficiently large m.*

This is the immediate consequence of the definition of ample sheaf and the generalization of Serre's results on projective schemes (III-2.2.2).

PROPOSITION 8. *Let $\mathcal{L}$ be an ample sheaf on a projective scheme X and let Y be any closed subscheme of X. Then $\mathcal{L} \otimes \mathcal{O}_Y$ is also an ample sheaf on Y.*

Proof. Let j be the injection $Y \to X$. Then $\mathcal{L} \otimes \mathcal{O}_Y = j^*(\mathcal{L})$ and $\mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y = j^*(\mathcal{L}^{\otimes -m})$. Hence if $\mathcal{L}^{\otimes -m}$ is very ample, $\mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y$ is also very ample by (II-4.4.10). Q.E.D.

Let $X = \mathrm{Proj}\,(S)$ be a projective schema over a field k and let $S = k[X_0, X_1, \cdots, X_r]/\mathfrak{A}$ where $\mathfrak{A}$ is a homogeneous ideal of $k[X]$. Let $\mathfrak{A} = \bigcap_{i=0}^{t} \mathfrak{Q}_i$ be a normal decomposition of $\mathfrak{A}$. Now assume that the radical $\sqrt{\mathfrak{Q}_0}$ of $\mathfrak{Q}_0$ is the irrelevant ideal $(X_0, X_1, \cdots, X_r)$. In this case let us put $\mathfrak{A}_1 = \bigcap_{i=1}^{t} \mathfrak{Q}_i$ and let $S_1 = k[X_0, X_1, \cdots, X_r]/\mathfrak{A}_1$. We shall show that X and $\mathrm{proj}\,(S_1) = X_1$ are homeomorphic. Let $\mathfrak{J}$ be the kernel of the homomorphism $\phi : S \to S_1$. Then we have an exact sequence

$$0 \to \mathfrak{J} \to S \xrightarrow{\phi} S_1 \to 0.$$

It is not difficult to see that $\mathfrak{J}$ is composed of elements x such that $xS_+^n = 0$ for some n. Hence any prime ideal $\mathfrak{p}$ in $\mathrm{proj}\,(S)$ must contain the ideal $\mathfrak{J}$ and there is a bijection from X onto X_1. The coincidence of the structure sheaves can be seen immediately and thus X and X_1 are homeomorphic. The above consideration shows that any projective scheme over a field k has a homogeneous coordinate ring S such that any generator of S contains at least one non-zero divisor. This implies, among others, the coherent sheaf $\mathcal{O}_X(-1) = \mathcal{O}_X(1)^{-1}$ can be represented by a divisor $\mathcal{H}$ called a *hyperplane divisor* of X.

PROPOSITION 9. *Let $\mathcal{L}$ be an ample sheaf and let $\mathcal{H}$ be a hyperplane divisor on X. Then there exists an integer N such that*

$$H^{\rho}(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{H}^{\otimes -s}) = \{0\}$$

if $m \geqq N$, $s \geqq 0$ and $\rho \geqq 1$.

Proof. We shall use the induction on the dimension n of X. Let H be a closed subscheme defined by a hyperplane divisor. Then we have an exact sequence of sheaves.

$$0 \to \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s+1} \to \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s} \to \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s} \otimes \mathcal{O}_H \to 0.$$

Since $\mathscr{L} \otimes \mathcal{O}_H$ is ample on H by Proposition 8, there exists an integer N_1 such that

$$H^\rho(X, \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s} \otimes \mathcal{O}_H) = H^\rho(H, \mathscr{L}_1{}^{\otimes -m} \otimes \mathscr{H}_1{}^{\otimes -s}) = \{0\}$$
$$\text{for } m \geqq N_1,\ s \geqq 0, \text{ and } \rho \geqq 1$$

by the induction assumption, where $\mathscr{L}_1 = \mathscr{L} \otimes \mathcal{O}_H$, $\mathscr{H}_1 = \mathscr{H} \otimes \mathcal{O}_H$. Let N_2 be an integer such that $H^\rho(X, \mathscr{L}^{\otimes -m}) = \{0\}$ whenever $m \geqq N_2$ and $\rho \geqq 1$. The integer $N = \mathrm{Max}(N_1, N_2)$ will satisfy the requirement of the proposition. In fact, if we fix an integer $m \geqq N$, we have the exact sequence of cohomology groups

$$H^\rho(X, \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s+1}) \to H^\rho(X, \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s}) \to 0.$$

for $\rho \geqq 1$. Beginning from $s = 1$, we see immediately that

$$H^\rho(X, \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s}) = \{0\}$$

for any $s \geqq 0$ and $\rho \geqq 1$. Q. E. D.

The similar device used above leads to the following

PROPOSITION 10. *Let $\mathscr{L}$, $\mathscr{H}$ be as in Proposition 8, let $\mathscr{D}$ be an arbitrary invertible sheaf on X. Then there exists an integer N such that*

$$H^\rho(X, \mathscr{L}^{\otimes -m} \otimes \mathscr{H}^{\otimes -s} \otimes \mathscr{D}) = \{0\}$$

if $m \geqq N$, $s \geqq 0$, and $\rho \geqq 1$.

5. Proof of Theorem 3 (the case $n > 1$). In this paragraph we shall assume that X is an irreducible variety over k embedded in a projective space. Let $\mathscr{L}$ be an arithmetically positive divisor on X. We shall show $\mathscr{L}$ is an ample divisor under the induction assumption in Theorem 3. For the sake of simplicity we shall denote by $h^\rho(m)$ the dimension (over k) of the cohomology group $H^\rho(X, \mathscr{L}^{\otimes -m})$. The proof will be divided in several steps. The assertion involving the integer m is claimed to hold only if m is sufficiently large, though sometimes it may not be mentioned explicitly.

(I) $$h^\rho(m) = 0 \ \text{ for } \rho \geqq 2.$$

Proof. Let $\mathfrak{H}$ be a hyperplane divisor. Then we have an exact sequence of coherent sheaves.

$$0 \to \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s+1} \to \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s} \to \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s} \otimes \mathcal{O}_H \to 0.$$

where H is the hyperplane section of X defined by a hyperplane divisor $\mathfrak{H}$. From this we have an exact sequence of cohomology groups

$$H^{\rho-1}(H, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s} \otimes \mathcal{O}_H) \to H^{\rho}(X, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s+1})$$
$$\to H^{\rho}(X, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s}) \to H^{\rho}(H, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s} \otimes \mathcal{O}_H).$$

The two extreme terms vanish by Proposition 10 if m is big enough and $\rho \geqq 2$. Hence we have an equality

$$\dim H^{\rho}(X, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s+1}) = \dim H^{\rho}(X, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s})$$

for any $s \geqq 0$, $\rho \geqq 2$, if m is big enough. For a fixed integer m we know that $H^{\rho}(X, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{H}^{\otimes -s}) = \{0\}$ if s is big enough, by (III-2.2.2), hence $H^{\rho}(X, \mathfrak{L}^{\otimes -m}) = \{0\}$ for any $\rho \geqq 2$.

The following proposition will be proved by the same device.

(I′) Let $\mathfrak{D}$ be an arbitrary divisor on X. Then $H^{\rho}(X, \mathfrak{L}^{\otimes -m} \otimes \mathfrak{D}) = \{0\}$ if $\rho \geqq 2$.

(II) Assume moreover that $\mathfrak{L}$ is a positive divisor. Then $h^1(m)$ is a constant after some large integer m.

Proof. Since $\mathfrak{L}$ is positive we have an exact sequence

$$0 \to \mathfrak{L}^{\otimes -m+1} \to \mathfrak{L}^{\otimes -m} \to \mathfrak{L}^{\otimes -m} \otimes \mathcal{O}_L \to 0,$$

where L is the closed subscheme defined by the sheaf $\mathfrak{L}$. Since $h^{\rho}(\mathfrak{L}^{\otimes -m} \otimes \mathcal{O}_L) = 0$ for large m and $\rho \geqq 1$, by induction assumption, we see that $h^1(m-1) \geqq h^1(m)$ after a certain integer. The inequality cannot hold indefinitely.

(III) Under the same assumption as in (II), let $\mathfrak{L} = \mathfrak{D} \otimes \mathfrak{D}'$ be a decomposition of $\mathfrak{L}$ as a product of two positive divisors $\mathfrak{D}$ and $\mathfrak{D}'$. Then $h^1(\mathfrak{L}^{\otimes -m} \otimes \mathfrak{D}) = h^1(m)$.

Proof. From the inclusion relation

$$\mathfrak{L}^{\otimes -m+1} \subset \mathfrak{L}^{\otimes -m+1} \otimes \mathfrak{D}'^{-1} = \mathfrak{L}^{\otimes -m} \otimes \mathfrak{D} \subset \mathfrak{L}^{\otimes -m}$$

and $h^1(\mathfrak{L}^{\otimes -m} \otimes \mathcal{O}/\mathfrak{D}) = 0$, $h^1(\mathfrak{L}^{\otimes -m+1} \otimes \mathfrak{D}'^{\otimes -1} \otimes \mathcal{O}/\mathfrak{D}') = 0$ by induction assumption, we get easily the relations

$$h^1(m-1) \geqq h^1(\mathfrak{L}^{\otimes -m+1} \otimes \mathfrak{D}') = h^1(\mathfrak{L}^{\otimes -m} \otimes \mathfrak{D}) \geqq h^1(m).$$

Q. E. D.

The inequality combined with (II), we get the assertion.

COROLLARY. *Retaining the notations and assumptions as in* (III), *we have an exact sequence*

$$0 \to H^0(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{D}) \to H^0(X, \mathcal{L}^{\otimes -m}) \to H^0(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{O}/\mathcal{D}) \to 0.$$

(IV) Let D be an arbitrary divisor on X. Then $H^0(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{D}) \neq \{0\}$.

By assumption,

$$\chi(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{D}) = \dim H^0(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{D}) - \dim H^1(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{D})$$

increases infinity as m increases, since $h^\rho(\mathcal{L}^{\otimes -m} \otimes \mathcal{D}) = 0$ for $\rho \geqq 2$. In particular, $H^0(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{D})$ is not zero.

(V) The morphism γ_m associated to the divisor $\mathcal{L}^{\otimes m}$ is regular everywhere on X.

Proof. It is sufficient to show that γ_m is regular at every point x algebraic over k. Since a divisor $\mathcal{L}$ is ample if and only if $\mathcal{L}^{\otimes m}$ is ample for some m (II-4. 6. 9), we can assume without loss of generality that $\mathcal{L}$ is a positive divisor. In this case 1 is a section of $H^0(X, \mathcal{L}^{\otimes -1})$ and hence any point outside the support of $\mathcal{L}$ is contained in the domain of regularity of the morphism γ_m. Let m be a large integer such that the sequence

$$(0) \to H^0(X, \mathcal{L}^{\otimes -m+1}) \to H^0(X, \mathcal{L}^{\otimes -m}) \to H^0(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{O}/\mathcal{L}) \to (0)$$

is exact. Let L be the closed subscheme defined by the sheaf $\mathcal{L}$. Since $H^0(X, \mathcal{L}^{\otimes -m} \otimes \mathcal{O}/\mathcal{L}) = H^0(L, \bar{\mathcal{L}}^{\otimes -m})$ where $\bar{\mathcal{L}} = \mathcal{L} \otimes \mathcal{O}_L$ and $\bar{\mathcal{L}}$ is an ample sheaf on L by induction assumption, there exists a section $\bar{s}$ in $H^0(L, \bar{\mathcal{L}}^{\otimes -m})$ such that $\bar{s}(\bar{x}) \neq 0$ for a given $\bar{x}$ in L, if m is large enough. Let s be a section of $H^0(X, \mathcal{L}^{\otimes -m})$ whose image in $H^0(L, \bar{\mathcal{L}}^{\otimes -m})$ is precisely $\bar{s}$. Then it is not difficult to see that $s(\bar{x}) \neq 0$, proving γ_m is regular also at $\bar{x}$.

(VI) $\gamma_m^*(\mathcal{O}_Y(1)) \cong \mathcal{L}^{\otimes -m}$.

The similar device used in Section 3 can be applied to this case. Hence the proof is omitted.

(VII) Let y be a point of Y rational over k and assume that $\mathcal{L} = \gamma_1^*(\mathcal{O}_Y(1))$. Then $\gamma_1^{-1}(y)$ contains only a finite number of points.

Proof. Let $\mathfrak{J}$ be a hyperplane divisor of Y not containing y. Since $\mathfrak{J}^{-1} \cong \mathcal{O}_Y(1)$, we have $\gamma^*(\mathfrak{J}^{\otimes -m}) = \mathcal{L}^{\otimes -m}(\gamma = \gamma_1)$. Assume that $\gamma^{-1}(y)$ contains infinitely many points, then it implies $\gamma^{-1}(y)$ is a closed subscheme of X of positive dimension. We shall denote this subscheme by D. The structure sheaf $\mathcal{O}_D$ is given by $\mathcal{O}_D = \mathcal{O}_X \otimes_{\mathcal{O}_Y}(\mathcal{O}_Y/\mathfrak{M})$ where $\mathfrak{M}$ is the defining sheaf of the closed point y. Since

$$\mathcal{L}^{\otimes -m} \otimes_{\mathcal{O}_X} \mathcal{O}_D = (\mathfrak{J}^{\otimes -m} \otimes_{\mathcal{O}_Y} \mathcal{O}_Y) \otimes_{\mathcal{O}_Y}(\mathcal{O}_Y/\mathfrak{M})$$

and $\mathfrak{J}_y = \mathcal{O}_y$, we have

$$(\mathcal{L}^{\otimes -m} \otimes_{\mathcal{O}_X} \mathcal{O}_D)_x = \begin{cases} \mathcal{O}_D & \text{if } x \in D \\ 0 & \text{if } x \notin D \end{cases}$$

and $\chi(\mathcal{L}^{\otimes -m} \otimes_{\mathcal{O}_X} \mathcal{O}_D) = \text{const.}$ for any m.

This contradicts the assumption that $\mathcal{L}$ is an arithmetically positive divisor.

(VIII) If the morphism $f: X \to Y$ satisfies the condition of (VII) for any rational point over k, then $f^{-1}(y)$ is also a finite set for any point y of Y.

Proof. By (III-4.4.1) the set of points $x \in X$, such that $f^{-1}(f(x))$ is a finite set forms an open subset X' of X. Hence if $X \neq X'$ there must exist a k rational point x in $X - X'$ such that $f^{-1}(f(x))$ is not a finite set, contradicting the assumption. Q.E.D.

The assertion (VIII) completes the proof of Theorem 3 on account of Theorem 1.

6. Ample divisors on a non-singular variety. In this paragraph we shall give an alternative condition for a divisor to be arithmetically positive. Let V^n be a non-singular variety of dimension n in a projective space and let $\mathfrak{G}_r(V)$ be the additive group of cycles of codimension r on V. Let $\mathfrak{G}_r^0(V)$ be the subgroup of $\mathfrak{G}_r(V)$ consisting of cycles which are numerically equivalent to zero. Let $\mathfrak{G}_r^*(V) = \mathfrak{G}_r(V)/\mathfrak{G}_r^0(V)$. Then the direct sum $\mathfrak{G}^*(V) = \sum_{r \geqq 0} \mathfrak{G}_r^*(V)$ ($\mathfrak{G}_r^*(V) = 0$ if $r > n$) will be a graded ring in which the multiplication is defined by the intersection product of cycles on V. Let x be a homogeneous element of $\mathfrak{G}^*(V)$. We shall say x is numerically positive if the class x contains a positive cycle.

Since $\mathfrak{G}_n(V)$ is isomorphic to the ring of rational integers Z it is meaningful to say that an element of degree n in $\mathfrak{G}^*(V)$ is positive, negative, or zero.

THEOREM 4. *Let X be a divisor on a non-singular variety V and let x be a class of X in $\mathfrak{G}^*(V)$. Assume that $(x^r y) > 0$ for any numerically positive element y of degree $n-r$ $(r=1,\cdots,n)$. Then X is an ample divisor, and vice versa.*

Proof. On a non-singular variety V any reduced scheme of dimension r corresponds, in a 1-1 way, to a cycle Y of dimension r whose reduced expression is the sum of irreducible varieties with coefficients 1. (We shall call such a cycle a reduced cycle). Hence to prove the theorem it suffices to prove the following

PROPOSITION 11. *Let X, V be as in Theorem 4 and let Y be a reduced cycle of dimension r. Let $\mathcal{L}$ be the locally free sheaf on V defined by X. Then $\chi(\mathcal{L}^{\otimes -m} \otimes O_Y)$ is a polynomial of degree r in m whose coefficient of highest term is given by $(x^r y)/r!$.*

Proof. We shall prove by using double induction on the dimension n of V and dimension r of Y.

The case $n=1$ is contained in the case $r=1$.

In the case where $r=1$, it is proved in Proposition 4.

Let D and E be non-singular subvarieties of codimension 1 such that $X+D \sim E$ and such that the cycles $Y \cdot D$, $Y \cdot E$ are also reduced cycles. Moreover, we can assume without loss of generalities every component of X, D, and E does not contain any component of Y. We shall denote by $\mathcal{L}$, $\mathcal{D}$ and $\mathcal{E}$ the invertible sheaves defined by X, D, and E, respectively. We shall consider exact sequences

$$0 \to \mathcal{L}^{\otimes -m} \otimes \mathcal{D} \otimes \mathcal{O}_Y \to \mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y \to \mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y \otimes \mathcal{O}_D \to 0$$

$$\begin{aligned} 0 \to \mathcal{L}^{\otimes -m-1} \otimes \mathcal{E} \otimes \mathcal{O}_Y \to {} & \mathcal{L}^{\otimes -m-1} \otimes \mathcal{O}_Y \\ & \to \mathcal{L}^{\otimes -m-1} \otimes \mathcal{O}_Y \otimes \mathcal{O}_E \to 0. \end{aligned}$$

Since $\mathcal{L}^{\otimes -m} \otimes \mathcal{D} = \mathcal{L}^{\otimes -m-1} \otimes \mathcal{E}$, we get

$$\begin{aligned} & \chi(\mathcal{L}^{\otimes -m-1} \otimes \mathcal{O}_Y) - \chi(\mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y) \\ & \qquad = \chi(\mathcal{L}^{\otimes -m-1} \otimes \mathcal{O}_Y \otimes \mathcal{O}_E) - \chi(\mathcal{L}^{\otimes -m} \otimes \mathcal{O}_Y \otimes \mathcal{O}_D). \end{aligned}$$

The right-hand side is, by induction assumption, a polynomial in m of degree $(r-1)$ and the coefficient of m^{r-1} is given by

$$[(x^{r-1}ye) - (x^{r-1}yd)]/(r-1)! = \frac{(x^r y)}{(r-1)!},$$

where e and d denote the classes of D and E in $\mathfrak{G}^*(V)$ respectively. From this we can draw easily the conclusion in a routine way. Q. E. D.

HIROSHIMA UNIVERSITY AND HARVARD UNIVERSITY.

REFERENCES.

[1] A. Grothendieck, "Éléments de Géométrie Algébrique, I, II, III," *Institut des Hautes Études Scientifiques* (1960-1961).

[2] Y. Nakai, "Non-degenerate divisors on an algebraic surface," *Journal of Science of the Hiroshima University*, vol. 24 (1960).

[3] ———, "Some fundamental lemmas on projective schemes," *to appear.*

[4] J. P. Serre, *Groupes algébriques et corps de classes*, Hermann, Paris.

PERIODIC SOLUTIONS OF THE RESTRICTED THREE BODY PROBLEM REPRESENTING ANALYTIC CONTINUATIONS OF KEPLERIAN ELLIPTIC MOTIONS.*

By RICHARD F. ARENSTORF.

Introduction. The equations of motion for the plane Restricted Three Body Problem can be written in the form

$$x'' + 2ix' - x$$
$$(1) \qquad = -(1-\mu)(x+\mu)\,|\,x+\mu\,|^{-3} - \mu(x+\mu-1)\,|\,x+\mu-1\,|^{-3}, \qquad (' = d/dt)$$

where $x = x_1 + ix_2$ is the complex position vector of the infinitesimal body referred to a co-system rotating with angular velocity 1 about the center of gravity of the two attracting bodies of masses $1-\mu$ and μ $(0 \leqq \mu \leqq 1)$ as origin.

When $\mu = 0$, the solutions of (1) are well known and can be represented as $x(t) = e^{-it}z(t)$, where the complex position vector $z(t)$ describes the Keplerian motion: i.e. a solution of $z'' = -z\,|\,z\,|^{-3}$. Under suitable initial conditions this latter motion will be periodic, for instance with

$$(2) \qquad z(0) = a(1+\epsilon), \quad z'(0) = ic^*/z(0), \quad c^{*2} = a(1-\epsilon^2), \qquad (a > 0, 0 < \epsilon < 1)$$

$z(t)$ moves along an ellipse with major half axis a and eccentricity ϵ, having $z = 0$ as focus and $z(0)$ at maximum distance from 0. Its sidereal period is $T_0 = 2\pi\,|\,a^{\frac{3}{2}}\,|$. The corresponding $x(t)$ will be periodic, iff. T_0 is commensurable with 2π, or $a^{\frac{3}{2}} = m/k$, where k and m are relatively prime integers, $m > 0$ and k is chosen positive resp. negative, if $z(t)$ is direct resp. retrograde; i.e., $\operatorname{sign} k = \operatorname{sign} c^*$. The synodical period on the rotating ellipse then is $T^* = 2\pi m$ and the curve $x = x(t)$, $(0 \leqq t \leqq T^*)$ is closed after $k - m$ positive revolutions around the origin. We denote this solution of (1) with $\mu = 0$ from now on by $x^*(t)$ and obtain from (2) for its initial values

$$(3) \qquad x^* = a(1+\epsilon), \quad dx^*/dt = i(c^* - x^{*2})/x^* \quad \text{at } t = 0.$$

We shall show the existence of periodic solutions $x(t)$ of (1) for small $\mu > 0$, which are near the generating solutions $x^*(t)$ belonging to arbitrary integers

* Received September 13, 1962.

$k, m \neq 0$ and properly restricted ϵ: Namely, there are for fixed $a = (m/k)^{2/3}$ at most finitely many ϵ in $0 < \epsilon < 1$ with $\epsilon = \sqrt{(1 - a^{-3})}$ or with $x^*(t) = 1$ at least once in $0 \leqq t \leqq T^*$. For every closed ϵ-interval I containing none of these exceptional values there exists a positive μ^* such that (1) possesses for every fixed μ in $0 \leqq \mu < \mu^* \leqq 1$ a family of periodic solutions depending analytically upon the parameter ϵ in I. These solutions are holomorphic also in μ and transfer into $x^*(t)$ for $\mu = 0$. Their synodical periods and Jacobi-constants are holomorphic in ϵ and μ and depend both actually upon μ.

This result includes especially the existence of the periodic solutions of the so-called second kind for the Restricted Three Body Problem. Their existence had been claimed with supposed proofs by H. Poincaré [5], K. Schwarzschild [6] and C. L. Charlier [2], whose invalidity was shown by P. Staeckel [8] and A. Wintner [9], [10], however. In these attempts the continuation method of Poincaré was employed in an isoperiodic or an isoenergetic manner. We too, apply this continuation method, but replace the general (because of the existence of the Jacobi integral unsymmetric) periodicity condition of Poincaré by a more special one, which is based on the symmetry of the dynamical problem (1), and has been used by G. D. Birkhoff [1] already. This condition is not only simpler and more natural, it also points out and reduces the redundancy in the classical periodicity conditions, which caused the critical functional determinants to vanish; see A. Wintner [9] and C. L. Siegel [7]. Otherwise, we achieve our goal by employing appropriate variables, which render the dependence of the Keplerian motion upon its initial values in a most simple form.

In this regard, it is to be mentioned that G. D. Birkhoff [1] showed a. o. the existence of periodic solutions of (1) for small μ, which close in the rotating co-system only after sufficiently many revolutions about the mass $1 - \mu$, if their Jacobi constant determines a simple closed zero-velocity curve containing the orbits in its interior. B. O. Koopman [3] established the analogon for the exterior case of a zero-velocity curve with forbidden bounded interior. More recently J. Moser [4] showed the existence of periodic solutions of (1) for small μ, which close after many revolutions and are near solutions of the existing first kind of Poincaré (generated from circular motions for $\mu = 0$). All these solutions correspond for $\mu = 0$ to periodic motions along rotating ellipses with rational $a^{\frac{3}{2}}$. It is presently not known if these solutions for $\mu > 0$ coincide with certain of our above solutions $x(t)$ generated from $x^*(t)$ with large $|k - m|$ and suitable, small $\epsilon > 0$.

Finally we remark that several of these solutions for different m/k, whose existence is shown here for small $\mu > 0$, have been numerically calculated by

us for increasing μ on high speed electronic computers. They are particularly of interest, when $a(1-\epsilon) < \delta$ and $1 < a(1+\epsilon) < 1+\delta$ with small $\delta > 0$, since then they pass repeatedly near both masses of the Restricted Three Body Problem. The calculations indicate their existence for values of μ at least as large as that for the case $1-\mu/\mu$ equal to mass of the Earth/mass of the Moon. Thus, their practical significance for astronautics is apparent. This also was one of the incentives for the investigation presented here.

Existence Proof. Let $x(t)$ be a solution of (1), which is holomorphic on an interval $0 \leqq t \leqq T$, $T > 0$; i.e. free of collisions. If (with a bar denoting the conjugate complex number)

$$x(\tfrac{1}{2}T) = \bar{x}(\tfrac{1}{2}T), \qquad x'(\tfrac{1}{2}T) = -\bar{x}'(\tfrac{1}{2}T), \tag{4}$$

then the function $\bar{x}(T-t)$ of t is identical with $x(t)$, since it satisfies (1) also and the two functions and their first derivatives coincide respectively at $t = \frac{1}{2}T$. This implies, if additionally

$$x(0) = \bar{x}(0), \qquad x'(0) = -\bar{x}'(0), \tag{5}$$

that $x(T) = x(0)$, $x'(T) = x'(0)$, so that $x(t)$ will be periodic with period T, since (1) is autonomous. Then the closed curve $x = x(t)$, $(0 \leqq t \leqq T)$ is symmetric over the x_1-axis, since $x(-t) = \bar{x}(t)$. Especially $x^*(t)$ satisfies (5) by (3), and also (4) with $T = T^* = |k|\,T_0$, since as a consequence of (2) $z(k \cdot T_0/2)$ is real and $z'(k \cdot T_0/2)$ is pure imaginary for every integer k.

We introduce new real variables F, H, U, V instead of $x = x_1 + ix_2$ and $y = x' + ix = y_1 + iy_2$ by

$$\left\{\begin{array}{ll} F = \operatorname{arctg} x_2/x_1,\ H = \frac{1}{2}(y_1^2 + y_2^2) - r^{-1} - c, & r = (x_1^2 + x_2^2)^{\frac{1}{2}}, \\ U = x_1/r - cy_2,\ \ V = x_2/r + cy_1, & c = x_1y_2 - x_2y_1. \end{array}\right. \tag{6}$$

The functional determinant of this transformation is $D = -(2H + 3c) \cdot c^2/r^2$. Now c and H are first integrals of (1) for $\mu = 0$, since with $z(t) = e^{it}x(t)$ then

$$c = \operatorname{Im} \bar{x}y = \operatorname{Im} \bar{z}z' = \text{const. of area},$$
$$H + c = \tfrac{1}{2}\,|z'|^2 - |z|^{-1} = \text{const. of energy},$$

as well known for the Keplerian motion. By (2) then

$$c = c^*,\ H = -c^* - 1/2a,\ \ r = |x^*(t)| \geqq a(1-\epsilon) > 0 \ \text{ on } x^*(t).$$

Thus

$$D = -(c^* - 1/a)\,c^{*2}/r^2 \ \text{ on } x^*(t), \tag{7}$$

and the transformation (6) is analytic and locally 1-to-1 in a neighborhood of every point on the trajectory $x^*(t), y^*(t) = dx^*(t)/dt + ix^*(t)$,

$(0 \leqq t \leqq T^*)$ if $D \neq 0$ or $ac^* \neq 1$. This holds for $0 < \epsilon < 1$ always, when $a \leqq 1$ or when $\epsilon \neq \sqrt{(1-a^{-3})}$ for $a > 1$, $c^* > 0$, and this assumption will be made from here on.

Now (6) transforms (1) in case $\mu = 0$ into

$$(8) \qquad \begin{cases} F' + 1 = c/r^2 = c^{-3}(1 - U\cos F - V\sin F)^2, \\ H' = 0,\ U' = V,\ V' = -U, \end{cases}$$

since from (6) $c = r(1 - U\cos F - V\sin F)/c$. If we denote initial values by the corresponding small letter, (8) can be integrated up to

$$(9) \qquad \begin{cases} H = h,\ U + iV = (u + iv)e^{-it}, \\ F' + 1 = c^{-3}(1 - u\cos(F+t) - v\sin(F+t))^2. \end{cases}$$

This and (6) yield

$$(10) \qquad u^2 + v^2 = U^2 + V^2 = 1 - 2c^2/r + c^2(y_1{}^2 + y_2{}^2) = 1 + 2hc^2 + 2c^3,$$

so that c depends upon the initial values h, u, v. Substituting for these the special values

$$(11) \qquad f^* = 0,\ h^* = -c^{*} - 1/2a,\ u^* = \epsilon,\ v^* = 0$$

derived from (3) and (6) gives $c = c^*$ and by (9) the original solution $x^*(t)$, but now represented in the new variables.

In general (6) transforms (1) into a system

$$(12) \qquad F' = g_1, \qquad H' = g_2, \qquad U' = g_3, \qquad V' = g_4,$$

where the $g_n = g_n(F, H, U, V, \mu)$ are holomorphic functions of all variables in a neighborhood of the special solution determined by (9), (11) and for sufficiently small $\mu \geqq 0$, if $x^*(t)$ is free of collisions for $\mu = 0$. Given $a = (m/k)^{2/3}$, such a collision, i. e. $x^*(t) = 1$ for some t in $0 \leqq t \leqq T^*$, can happen only for finitely many values of ϵ in $0 < \epsilon < 1$. We return to this condition later and assume here only that these ϵ are omitted. Then, according to Poincaré's extension of Cauchy's existence theorem for ordinary differential equations (for a modern proof see C. L. Siegel [7]) the solutions of (12) are holomorphic functions of t, f, h, u, v, and μ for $0 \leqq t \leqq 2T^*$, say, and sufficiently small $|f| + |h - h^*| + |u - \epsilon| + |v| + \mu$, $\mu \geqq 0$, using (11). For $\mu = 0$ (12) becomes (8).

We consider now the solutions of (12) with small $\mu > 0$ and initial values near (11). These solutions can be assumed remaining near the original solution given by (9), (11) for $0 \leqq t \leqq 2T^*$, so that especially (6) is

applicable. In order that such a solution will be periodic in the former coordinates x_1, x_2 with period $T > 0$, it is sufficient by (4), (5), (6) that

$$(13) \qquad \begin{cases} F(\tfrac{1}{2}T, f, h, u, v, \mu) = \pi(k - m), & f = 0, \\ V(\tfrac{1}{2}T, f, h, u, v, \mu) = 0, & v = 0, \end{cases}$$

since for the considered solutions of (12) c remains near c^*, and therefore the assumption $c \neq 0$ at $t = \frac{1}{2}T$ can be made, if $T < 2T^*$. These equations are actually satified for the original solution $x^*(t)$ and $\mu = 0$, or with (11) for

$$(14) \qquad T = T^* = 2m\pi,\ f = f^*,\ h = h^*,\ u = u^*,\ v = v^*,\ \mu = 0,$$

since

$$\operatorname{arc} x^*(\tfrac{1}{2}T^*) = \operatorname{arc}[e^{-im\pi}z(|k|T_0/2)] = \operatorname{arc}[e^{-im\pi}e^{ik\pi}] = (k - m)\pi$$

by (2). Hence (13) can be satisfied for small $\mu > 0$ and initial values near (14), if for instance the functional determinant

$$(15) \qquad D^* = F_t V_h - F_h V_t \neq 0 \text{ for } t = \tfrac{1}{2}T \text{ at } (14).$$

By the holomorphy it suffices putting $\mu = 0$ in F, V first and then calculate the partial derivatives, which therefore can be found from (9) and (10). Denoting partial derivatives by a corresponding index, (9) gives $V_h = 0$ and $V_t = -U$, hence with (11) and (14)

$$(16) \qquad D^* = \epsilon \cos m\pi \cdot F_h(m\pi, 0, h^*, \epsilon, 0), \quad (\mu = 0).$$

We put

$$(17) \qquad \phi = t + F(t, 0, h, \epsilon, 0) = \phi(t, h).$$

Then it follows from (9) that ϕ is uniquely determined by inversion of the integral

$$(18) \qquad t = c^3 \int_0^\phi (1 - \epsilon \cos\psi)^{-2} d\psi.$$

The value of c here is determined by $\epsilon^2 = 1 + 2hc^2 + 2c^3$ from (10). This implies

$$(19) \qquad c_h = -c(2h + 3c)^{-1}.$$

Differentiating (18) partially with respect to h gives

$$0 = 3c^2 c_h \int_0^\phi (1 - \epsilon \cos\psi)^{-2} d\psi + c^3 (1 - \epsilon \cos\phi)^{-2} \phi_h,$$

hence with (19)

$$(20)\quad \phi_h(m\pi, h^*) = 3(1-\epsilon\cos\phi^*)^2(2h^*+3c^*)^{-1}\int_0^{\phi^*}(1-\epsilon\cos\psi)^{-2}d\psi,$$

where $\phi^* = \phi(m\pi, h^*)$. But

$$2c^{*3}\int_0^{\pi}(1-\epsilon\cos\psi)^{-2}d\psi = \pm T_0 = 2\pi m/k$$

for the period on the original Keplerian ellipse as well known, thus by (18) $\phi^* = k\pi$ and by (16), (17), (20) finally

$$(21)\qquad D^* = 3\epsilon(-1)^m(1-\epsilon(-1)^k)^2 m\pi/(c^*-a^{-1})c^{*3} \neq 0,$$

so that (15) actually holds, since $0<\epsilon<1$ and $c^* \neq a^{-1}$ as required for (7).

From the implicit function theorem it follows now that (13) can be solved for T and h in a neighborhood of (14), and that $T-T^*$ and $h-h^*$ for $u=u^*=\epsilon$ result as power series in μ without constant terms having positive radii of convergence. This implies especially $0<T<2T^*$ for small $\mu>0$ as assumed before, so that by (13) the existence of the desired periodic solutions $x(t)$ is now actually shown. Their initial values $x(0)$, $x'(0)$ are determined by (6) and

$$(22)\qquad f=0,\ h=h(\mu,\epsilon),\ u=\epsilon,\ v=0$$

as functions of μ, ϵ and a.

Since for any given $a=(m/k)^{2/3}$ the foregoing is valid as long as ϵ is not one of the previously excepted values, it is a consequence of the local existence theorem for implicit functions and of the covering theorem that $T=T(\mu,\epsilon)$ and $h=h(\mu,\epsilon)$ are holomorphic functions of μ and ϵ on every closed ϵ-interval not containing one of the exceptional values and on $0\leqq\mu<\mu^*$ with corresponding sufficiently small $\mu^*>0$. The solutions of (12) belonging to (22) are holomorphic functions of their initial values and thus the corresponding periodic solutions $x(t)$ are holomorphic on $0\leqq t\leqq T$ and in ϵ and μ as stated in the introduction.

If one introduces in the transformation (6) instead of the variable H the Jacobi integral

$$J=\tfrac{1}{2}|y|^2-c-(1-\mu)|x+\mu|^{-1}-\mu|x+\mu-1|^{-1},$$

our whole consideration can be carried through in the same way. Thus also $J=J(\mu,\epsilon)$ is a holomorphic function of both variables, which clearly follows from

$$J=h+|x_0|^{-1}-(1-\mu)|x_0+\mu|^{-1}-\mu|x_0+\mu-1|^{-1},\quad x_0=x(0)$$

and the foregoing too. In (13) then h is to be replaced by the initial value j of J, but $J = j$ now. It is then of interest to consider besides (15) the other two functional determinants suggested by (13); namely,

$$D_1^* = F_t V_u - F_u V_t \quad \text{and} \quad D_2^* = F_j V_u - F_u V_j$$

for $t = \frac{1}{2}T$ at (14), where $j = h^*$. Since here $V_u = -\sin m\pi = 0$, $V_j = 0$ and $V_t = -U$ by (9), it follows $D_2^* = 0$ and $D_1^* = \epsilon \cos m\pi \cdot F_u(m\pi, 0, h^*, \epsilon, 0)$. Putting

$$\psi = t + F(t, 0, h^*, u, 0) = \psi(t, u),$$

it follows from (9) $\psi_t = c^{-3}(1 - u \cos \psi)^2$ and thus similar as from (18) for ϕ now $\psi(m\pi, u) = k\pi$ identically in $0 < u < 1$. Hence also $D_1^* = 0$. Thus solvability of (13) with respect to T, u for fixed $j = h^*$, or with respect to j, u for fixed $T = T^*$ remains at least doubtful, if at all possible. In fact, A. Wintner [9] has shown that for sufficiently small $\epsilon > 0$ isoperiodic solutions do not exist.

It is interesting to note that the treatment presented here is also effective for the periodic solutions of the first kind. In this case $\epsilon = 0$, and $a^{\frac{3}{2}} = \pm T_0/2\pi = \omega^{-1}$ can be taken arbitrary, especially not rational, with the exceptions $\omega \neq -1, 1$ and $\omega \neq 1 \pm m^{-1}$, m natural. Then it follows $D_1^* \neq 0$, $D_2^* \neq 0$, but $D^* = 0$ in (15), which is the reverse of our present situation for $\epsilon > 0$. Of course, an existence proof for the solutions of the first kind is long known.

Finally we consider the restrictions placed upon ϵ for our existence proof. These are

$$\begin{aligned} &0 < \epsilon < 1, \ \epsilon \neq \sqrt{(1 - a^{-3})} \ \text{ for } a > 1, \ c^* > 0, \\ (23) \qquad &x^*(t) \neq 1 \text{ in } 0 \leqq t \leqq m\pi \ \text{ for } a = (m/k)^{2/3}. \end{aligned}$$

The first of them is equivalent to $ac^* \neq 1$ or $2h^* + 3c^* \neq 0$ and was required for (7) and (21). The dynamical meaning of this condition becomes clear, when the sidereal frequency $\omega = 2\pi/T_0 = |a^{-\frac{3}{2}}|$ is used instead of a to characterize the generating elliptic motion together with its eccentricity. Since here $\mu = 0$, H coincides with the Jacobi-integral J, and then (19) shows that the Jacobi-constant $j = h$ of the generating motion $x^*(t)$, for which now

$$j^* = -c^* - 1/2a = -\tfrac{1}{2}\omega^{2/3} - \operatorname{sign} c^* \omega^{-1/3} \sqrt{(1 - \epsilon^2)}, \quad (\omega > 0)$$

has for given ϵ as function of c^* or ω a relative extremum (absolute maximum) for $ac^* = 1$ or $\omega = \sqrt{(1 - \epsilon^2)}$, in the direct case only. The inverse function

and thus the sidereal period T_0 then are not determined uniquely as functions of the Jacobi-constant and the eccentricity in a neighborhood of the branch points $(j^* = -(1-\epsilon^2)^{2/3}\cdot 3/2, \epsilon)$. The dynamical meaning of (23) is to exclude collisions with the perturbing body.

We shall show that for fixed k and m (23) excludes at most finitely many values of ϵ. To see this, we represent the Keplerian motion $z = z(t) = e^{it}x^*(t)$ with the help of the eccentric anomaly w, namely with (2) in the well-known form

$$(24) \qquad \begin{cases} z = a(\epsilon + \cos w + i\sqrt{(1-\epsilon^2)}\sin w), \\ t = a^{\frac{3}{2}}(w + \epsilon \sin w). \end{cases}$$

If (23) does not hold, then with appropriate ϵ between 0 and 1

$$(25) \qquad z = e^{it}, \qquad |z| = a(1 + \epsilon\cos w) = 1$$

for some w between 0 and $k\pi$ inclusively, using (24). The last equation implies $|1 - 1/a| \leqq \epsilon < 1$, thus $a > \frac{1}{2}$. Hence (23) holds always, when $2a \leqq 1$. If $a = 1$, then by (25) $\cos w = 0$ and thus with (24)

$$\epsilon = \cos t = \cos(w \pm \epsilon) = -\sin\epsilon,$$

which is impossible for $0 < \epsilon < 1$. Let now $a > \frac{1}{2}$, $a \neq 1$ and (25) be satisfied for appropriate ϵ and w. Then with $-A = 1 - 1/a$ and (24)

$$\epsilon\cos w = A \neq 0, \; a(\epsilon + \cos w) = \cos t,$$
$$(26) \qquad a(\cos w + A/\cos w) - \cos[a^{\frac{3}{2}}(w + A\operatorname{tg} w)] = 0.$$

Since $|A| \leqq \epsilon < 1$ and with w^* from $\cos w^* = |A|$, $0 \leqq w^* < \pi/2$, we have for suitable n

$$(27) \qquad |w - n\pi| < w^*, \quad n = \text{integer}, \quad |n| \leqq |k|.$$

Now the left side in (26) is a holomorphic function of w in (27) and even in every of the finitely many closed circles $|w - n\pi| \leqq w^*$. Thus it has at most finitely many zeros in (27), and these correspond to at most finitely many $\epsilon = A/\cos w$ satisfying (25). This proves our statement about (23).

One can easily see that our whole derivation remains valid, if we begin in (2) with $z(0) = a(1-\epsilon)$ at minimum distance from 0. This merely replaces ϵ by $-\epsilon$ in the subsequent equations. But it leads to new periodic solutions of (1) for small μ, which are actually different from the previous ones, even when considering solution curves only, if $k - m$ is odd. All of our periodic solution curves are symmetric over the x_1-axis, and they can be readily visualized in the rotating co-system.

NASA George C. Marshall Space Flight Center
and University of Alabama.

REFERENCES.

[1] G. D. Birkhoff, "The restricted problem of three bodies," *Rendiconti del Circolo Matematico di Palermo*, vol. 39 (1915), pp. 265-334.

[2] C. L. Charlier, *Die Mechanik des Himmels*, Vol. II, Chap. 9, para. 11, Walter De Gruyter and Co., Berlin (1927).

[3] B. O. Koopmann, "On rejection to infinity and exterior motion in the restricted problem of three bodies," *Transactions of the American Mathematical Society*, vol. 29 (1927), pp. 287-331.

[4] J. Moser, "Periodische Loesungen d. restringierten Dreikoerperproblems, die sich erst nach vielen Umlaufen schliessen," *Mathematische Annalen*, vol. 126 (1953), pp. 325-335.

[5] H. Poincaré, *Les Méthodes nouvelles de la mécanique céleste*, vol. I (1892), vol. III (1899).

[6] K. Schwarzschild, "Ueber eine Classe periodischer Loesungen des Dreikoerper-problems," *Astronomische Nachrichten*, vol. 147 (1896), pp. 17-24.

[7] C. L. Siegel, *Vorlesungen ueber Himmelsmechanik*. Springer Verlag, Berlin (1956), p. 156 and pp. 118-119.

[8] P. Stäckel, *Jahresberichte der Deutsche mathematiker-vereinigung*. 28 (1919), pp. 180-181.

[9] A. Wintner, "Ueber eine Revision der Sortentheorie des restringierten Dreikörper-problems," *Sitzungsberichte der Sächsischen Akademie der Wissenschaften zu Leipzig*, vol. 82 (1930), pp. 3-56.

[10] ———, "Grundlagen einer Genealogie der periodischen Bahnen im restringierten Dreikoerperproblem," *Mathematische Zeitscrift*, vol. 34 (1931), pp. 350-402.

TRANSITIVITY IN THE SPINORIAL KERNEL AND THE COMMUTATOR SUBGROUP OF THE ORTHOGONAL GROUP.*

By Barth Pollak.

Let V be a finite-dimensional vector space over a (commutative) field K of characteristic $\neq 2$ upon which is defined a quadratic form whose associated bilinear form is non-degenerate; that is, V is a non-singular quadratic space. (For the theory of quadratic spaces, see [1], [3], or [4].) Let $O(V)$ denote the orthogonal group of V, $O^+(V)$ the subgroup of elements having determinant $+1$, $O'(V)$ the subgroup of $O^+(V)$ consisting of the elements of spinor norm 1, and $\Omega(V)$ the commutator subgroup of $O(V)$. If $A \in V$ and $l(A)$ denotes the set of all vectors in V having the same length as A, then, by Witt's Theorem, $O(V)$ operates transitively on $l(A)$. If H is a subgroup of $O(V)$, H partitions $l(A)$ into a certain number $n(H)$ of transitivity classes. It is easy to compute $n(O^+(V))$. After certain preliminary results (§§ 1-2), we give (in § 3) an expression for $n(O'(V))$ in terms of spinor norms. If K is a field for which a representation theory for quadratic forms is known (i. e., finite fields, real numbers, local fields, global fields) we obtain (§§ 4-7) an explicit formula for $n(O'(V))$. Under a special assumption we are able to express $n(\Omega(V))$ as $n(O'(V))(O'(V):\Omega(V))$ which enables us to give an explicit formula for $n(\Omega(V))$ for the fields already mentioned. As a corollary we obtain necessary and sufficient conditions for $O'(V)$ and $\Omega(V)$ to operate transitively on $l(A)$.

For our convenience we adopt the notation of [1]. Thus if U is a subspace of V we denote its orthogonal complement by U^*, its dual space by U'. AB denotes the scalar product of $A, B \in V$ and $A^2 = AA$ denotes the *length* of A. The subspace spanned by vectors $X_1, \cdots, X_m$ will be denoted $\langle X_1, \cdots, X_m \rangle$. If $\langle A \rangle$ is non-singular, we denote the symmetry with respect to the hyperplane $\langle A \rangle^*$ by τ_A. If $\sigma \in O(V)$ we denote its spinor norm by $\theta(\sigma)$. Finally, $d(V)$ denotes the discriminant of V.

1. **Group-theoretic preliminaries.** Let S be a set and suppose G is a group that acts transitively on S. For $x \in S$, let G_x be the stability group of x. It is obvious that the points of S are in one-one correspondence with the elements of the coset space G/G_x. Now let H be a normal subgroup of G.

* Received September 14, 1962.

If $x, y \in S$ call $x \equiv y \bmod H$ if there exists $\sigma \in H$ such that $\sigma(x) = y$. This is an equivalence relation and partitions S into equivalence classes. Denote the set of equivalence classes by $S \bmod H$. Then G/H operates transitively on $S \bmod H$ in an obvious manner and if $\#(T)$ denotes the cardinality of the set T, we have

PROPOSITION 1. $\#(S \bmod H) = (G/H : G_x H/H)$ *for any* $x \in S$.

Proof. $G_x H/H$ is clearly the stability group of the equivalence class of $S \bmod H$ that contains x.

PROPOSITION 2. *Suppose N is a normal subgroup of G and $N \subseteqq H$. Then* $\#(S \bmod N) = (G/H : G_x H/H)(H/N : H_x N/N)$ *for any* $x \in S$.

Proof. This is nothing but a well-known index reduction principle applied to the canonical map of G/N onto G/H. (See [2], Lemma on p. 140.)

Remark 1. $N \subseteqq H$ implies $S \bmod N$ is a refinement of the partition $S \bmod H$. H/N operates transitively on an element of $S \bmod H$ and if $H(x)$ denotes that element of $S \bmod H$ containing x, then Proposition 2 could be restated (and reproved) as

$$\#(S \bmod N) = \#(S \bmod H)\,\#(H(x) \bmod N).$$

In the sequel we shall apply these propositions when $S =$ set of vectors of a given length and $G = O(V)$.

2. The subgroup $O_A(V)$. Set $O_A(V) = \{\sigma \in O(V) \mid \sigma(A) = A\}$ and $O_A^+(V) = O_A(V) \cap O^+(V)$. If $A^2 \neq 0$ then, trivially, $O_A(V) \cong O(\langle A \rangle^*)$. If A is non-zero and isotropic ($A^2 = 0$) we have

THEOREM 1. *Let V be a non-singular quadratic space of dimension $\geqq 3$, and suppose A is a non-zero isotropic vector. We may write $\langle A \rangle^* = \langle A \rangle \perp W$ with W non-singular. Then there exists a natural mapping $f : O_A(V) \to O(W)$ such that*

1) *f is an epimorphism*

2) $\operatorname{Ker} f \cong W'$

3) $\det \sigma = \det f(\sigma)$ *for each* $\sigma \in O_A(V)$

4) $\theta(\sigma) = \theta(f(\sigma))$ *for each* $\sigma \in O_A(V)$.

Proof. We may write $V = \langle A, B\rangle \perp W$ where $AB = 1$, $B^2 = 0$. If $C \in W$ we have for each $\sigma \in O_A(V)$, the unique decomposition

$$\sigma(C) = a_\sigma(C)A + b_\sigma(C)B + \sigma_W(C)$$

where $a_\sigma(C), b_\sigma(C) \in K$ and $\sigma_W(C) \in W$. And $0 = A \cdot C = A \cdot \sigma(C) = b_\sigma(C)$. A trivial computation shows that $a_\sigma \in W'$ and $\sigma_W \in \mathrm{Hom}_K(W, W)$. Also $C^2 = \sigma(C)^2 = (a_\sigma(C)A + \sigma_W(C))^2 = \sigma_W(C)^2$ hence $\sigma_W \in O(W)$ since W is non-singular. It is also easy to verify that for any $\sigma, \tau \in O_A(V)$, $(\sigma\tau)_W = \sigma_W\tau_W$. Hence the mapping $f\colon O_A(V) \to O(W)$ given by $f(\sigma) = \sigma_W$ is a morphism. If $\lambda \in O(W)$, set $\sigma = 1_{\langle A,B\rangle} \perp \lambda$ and we see that $f(\sigma) = \lambda$ whence f is an epimorphism. Now $\sigma \in \mathrm{Ker} f$ if and only if $\sigma_W = 1_W$. Thus $\sigma \in \mathrm{Ker} f$ if and only if $\sigma(C) = a_\sigma(C)A + C$ for all $C \in W$. Define a map $g\colon \mathrm{Ker} f \to W'$ by setting $g(\sigma) = a_\sigma$ and one readily verifies that g is an isomorphism. To compute $\det \sigma_W$ we write $\sigma(B) = xA + yB + D$ for some $D \in W$ and note that $1 = AB = A \cdot \sigma(B) = y$. Hence for a given choice of basis for W the matrix of σ, denoted (σ), has the shape

$$\begin{pmatrix} 1 & 0 & & \cdots & 0 \\ * & 1 & * & \cdots & * \\ * & 0 & & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & (\sigma_W) & \\ \cdot & \cdot & & & \\ * & 0 & & & \end{pmatrix}$$

hence $\det \sigma = \det f(\sigma)$. Finally we look at the spinor norm. Let $\tau_E \in O_A(V)$. Then $EA = 0$ and $E = xA + D$ for some $x \in K$, $D \in W$ and $0 \neq E^2 = D^2$. Thus if $C \in W$, $\tau_E(C) = -2x(CD/D^2)A + \tau_D(C)$ as a simple computation shows. Thus $f(\tau_E) = \tau_D$ and $\theta(\tau_E) = E^2K^{*2} = D^2K^{*2} = \theta(\tau_D) = \theta(f(\tau_E))$. Now let $\sigma_W \in O(W)$ and write $\sigma_W = \tau_{D_1} \cdots \tau_{D_r}$ with $D_i \in W$ for $i = 1, \cdots, r$. Set $\sigma = \tau_{E_1} \cdots \tau_{E_r}$ where $E_i = x_iA + D_i$ for some $x_i \in K$ $(i = 1, \cdots, r)$. Then $f(\sigma) = \sigma_W$ and $\theta(\sigma) = \theta(f(\sigma))$. Since every element of $O_A(V)$ whose image under f is σ_W is of the form $\sigma\lambda$ for some $\lambda \in \mathrm{Ker} f$ we will be through if we prove $\theta(\mathrm{Ker} f) = K^{*2}$. If $\lambda \in \mathrm{Ker} f$ we have seen that $\lambda(C) = a_\sigma(C)A + C$ for some $a_\sigma \in W'$. Set $W = \langle C_1\rangle \perp \cdots \perp \langle C_m\rangle$, $\rho_{C_i} = \tau_{-\frac{1}{2}a_\lambda(C_i)A - C_i}\tau_{C_i}$ for $i = 1, \cdots, m$ and $\rho = \rho_{C_1} \cdots \rho_{C_m}$. Then ρ and ρ_{C_i} $(i = 1, \cdots, m)$ are in $O^+(V)$ and we have $\rho_{C_i}(A) = A$, $\rho_{C_i}(C_i) = \lambda(C_i)$ for $i = 1, \cdots, m$ and if $i \neq j$, $\rho_{C_i}(C_j) = C_j$ for $j = 1, \cdots, m$. Hence $\rho(A) = A = \lambda(A)$ and $\rho(C_i) = \lambda(C_i)$ for $i = 1, \cdots, m$. Thus ρ and λ agree on the singular hyperplane $\langle A\rangle^*$ and hence agree on V (see [1], Theorem 3.17). Thus $\rho = \lambda$,

$$\theta(\lambda) = \prod_{i=1}^{m} \theta(\rho_{\sigma_i}) = K^{*2}$$

and the theorem is proved.

COROLLARY 1. $\theta(O_A^+(V)) = \theta(O^+(W))$.

Proof. Immediate by 3) and 4).

COROLLARY 2. *If* $\dim V = 3$ *then* $O_A^+(V) \cong K^+$, *the additive group of scalars.*

Proof. $\dim W = 1$ hence $O^+(W)$ is trivial and $W' \cong K^+$. (See [1], p. 133.)

For future convenience we set

$$l(A) = \{X \in V \mid X^2 = A^2, A \in V\}.$$

Then $O(V)$ operates transitively on $l(A)$. A subgroup H of $O(V)$ is called *length-transitive* if H operates transitively on $l(A)$ for each $A \in V$.

3. The basic result.

LEMMA 1. *Let V be a non-singular quadratic space of dimension $\geqq 2$ over K. If $A \in V$, then $\#(l(A) \bmod O^+(V)) = 1$ unless A is a non-zero isotropic vector and V is a hyperbolic plane in which case $\#(l(A) \bmod O^+(V)) = 2$.*

Proof. Of course $O(V)/O^+(V)$ is a cyclic group of order 2. If A is anisotropic, then $O_A(V)O^+(V)/O^+(V) \cong O_A(V)/O_A^+(V)$. There is a canonical isomorphism of $O_A(V)$ onto $O(\langle A\rangle^*)$ which maps $O_A^+(V)$ onto $O^+(\langle A\rangle^*)$. Thus $O_A(V)O^+(V)/O^+(V) \cong O(\langle A\rangle^*)/O^+(\langle A\rangle^*)$, a cyclic group of order 2. Now invoke Proposition 1. If $\dim V \geqq 3$ and A is non-zero isotropic then

$$O_A(V)O^+(V)/O^+(V) \cong O(W)/O^+(W) \text{ where } \langle A\rangle^* = \langle A\rangle \perp W,$$

W non-singular as one easily sees by Theorem 1. As before, invoke Proposition 1. Finally suppose V is a hyperbolic plane and A is non-zero isotropic. Then $O_A(V)$ is trivial. (See [1], Theorem 3.17.) By Proposition 1. $\#(l(A) \bmod O^+(V)) = 2$. q.e.d.

For future use we collect several well known facts and state them as

LEMMA 2. *Let V be a non-singular quadratic space. If* $\dim V \geqq 2$ *and V is isotropic (i.e., contains a non-zero isotropic vector), then* $\theta(O^+(V)) = K^*$. *If V is anisotropic, then* $\theta(O^+(V)) = K^{*2}$ *if* $\dim V = 1$; $N_{E/K}(E^*)$ *if* $\dim V = 2$ *where* $E = K((-d(V))^{\frac{1}{2}})$.

THEOREM 2. *Let V be a non-singular quadratic space of dimension* $\geqq 2$. *If A is any anisotropic vector, then*

$$\#(l(A) \bmod O'(V)) = (\theta(O^+(V)) : \theta(O^+(\langle A\rangle^*))).$$

If A is a non-zero isotropic vector, $\dim V \geqq 3$ *and* $\langle A\rangle^* = \langle A\rangle \perp W$, *$W$ non-singular, then*

$$\#(l(A) \bmod O'(V)) = (\theta(O^+(V)) : \theta(O^+(W))).$$

If A is a non-zero isotropic vector and $\dim V = 2$, *then*

$$\#(l(A) \bmod O'(V)) = 2(K^* : K^{*2}).$$

If V is isotropic then $\Omega(V) = O'(V)$ *and hence we obtain* $\#(l(A) \bmod \Omega(V))$.

If V is anisotropic and $O'(\langle A\rangle^*) = \Omega(\langle A\rangle^*)$ *then*

$$\#(l(A) \bmod \Omega(V)) = \#(l(A) \bmod O'(V))(O'(V) : \Omega(V)).$$

Proof. If A is anisotropic, $O_A^+(V) \cong O^+(\langle A\rangle^*)$ whence $\theta(O_A^+(V)) = \theta(O^+(\langle A\rangle^*))$. If A is non-zero isotropic and $\dim V \geqq 3$ then $\theta(O_A^+(V)) = \theta(O^+(W))$ by Corollary 1. Now $O^+(V)/O'(V) \cong \theta(O^+(V))$ and under this isomorphism $O_A^+(V)O'(V)/O'(V)$ is mapped onto $\theta(O_A^+(V))$. Now apply Proposition 2 to $S = l(A)$, $G = O(V)$, $H = O^+(V)$, $N = O'(V)$ and use Lemma 1. If A is non-zero isotropic and $\dim V = 2$ then apply Proposition 2 as before using the results of Lemma 1 and Lemma 2. That $O'(V) = \Omega(V)$ when V is isotropic is well known. (See [1], Theorem 5.17.). To finish the proof, we apply Proposition 2 with $S = l(A)$, $G = O^+(V)$, $H = O'(V)$ and $N = \Omega(V)$ and we see that all we need do is show that $(O'(V) \cap O_A^+(V))\Omega(V) = \Omega(V)$. Now

$$(O'(V) \cap O_A^+(V))\Omega(V)/\Omega(V) \cong O'(V) \cap O_A^+(V)/\Omega(V) \cap O_A^+(V).$$

There is a canonical isomorphism f mapping $O_A^+(V)$ onto $O^+(\langle A\rangle^*)$. f sends $O'(V) \cap O_A^+(V)$ onto $O'(\langle A\rangle^*)$ and $\Omega(V) \cap O_A^+(V)$ onto a subgroup H of $O^+(\langle A\rangle^*)$ containing $\Omega(\langle A\rangle^*)$. Hence

$$O'(V) \cap O_A^+(V)/\Omega(V) \cap O_A^+(V) \cong O'(\langle A\rangle^*)/H.$$

But $O'(\langle A\rangle^*) = \Omega(\langle A\rangle^*)$ by hypothesis. Thus $O'(\langle A\rangle^*)/H$ is trivial and the proof is complete.

COROLLARY 3. *Let V be a non-singular 2-dimensional quadratic space. Then $O'(V) = \Omega(V)$. If A is anisotropic, then*

$$\#(l(A) \bmod \Omega(V)) = (K^*: K^{*2})$$

if V is isotropic; $(N_{E/K}(E^*): K^{*2})$ if V is anisotropic where $E = K((-d(V))^{\frac{1}{2}})$. If A is non-zero isotropic, then $\#(l(A) \bmod \Omega(V)) = 2(K^*: K^{*2})$.

Proof. That $\Omega(V) = O'(V)$ for 2-dimensional non-singular quadratic spaces is well known. (See [1], Theorem 5.14.) The result follows immediately from Theorem 2 and Lemma 2.

4. K is a finite field.

LEMMA 3. *Let V be a non-singular quadratic space over a finite field K. Then $O'(V) = \Omega(V)$ and $\theta(O^+(V)) = K^{*2}$ if $\dim V = 1$; K^* if $\dim V \geqq 2$.*

Proof. If $\dim V \geqq 3$, then it is well known that V possesses non-zero isotropic vectors. Since $O'(V) = \Omega(V)$ for any K if $\dim V \leqq 3$, we have $O'(V) = \Omega(V)$ always. The assertions concerning $\theta(O^+(V))$ are also immediate with the exception of an anisotropic 2-dimensional space. But $K^* = N_{E/K}(E^*)$ is well known (see [2], Lemma on p. 131) and we are through.

THEOREM 3. *Let V be a non-singular quadratic space of dimension $\geqq 2$ over a finite field. Then $\Omega(V) = O'(V)$ and $\#(l(A) \bmod \Omega(V)) = 1$ if $\dim V \geqq 4$ or $\dim V = 3$ and A is anisotropic; 2 if $\dim V = 3$ and A is non-zero isotropic or $\dim V = 2$ and A is anisotropic; 4 if $\dim V = 2$ and A is non-zero isotropic.*

Proof. Immediate by Theorem 2, Corollary 3, Lemma 3 and the well known index relation $(K^*: K^{*2}) = 2$ for K a finite field of odd characteristic.

COROLLARY 4. *$O'(V) = \Omega(V)$ is length-transitive if and only if $\dim V \geqq 4$.*

5. K is the field of real numbers.

LEMMA 4. *Let V be a non-singular quadratic space over the field of real numbers. Then $\Omega(V) = O'(V)$ and $\theta(O^+(V)) = K^{*2}$ if V is anisotropic; K^* if V is isotropic.*

Proof. It is well known that $\Omega(V) = O'(V)$. (For a proof, use Theorem 5.14 and Theorem 5.16 of [1].) The assertion concerning $\theta(O^+(V))$ is obvious.

THEOREM 4. *Let V be a non-singular quadratic space of dimension $\geqq 2$ over the real numbers. Then $\Omega(V) = O'(V)$. If V is anisotropic, $\#(l(A) \bmod \Omega(V)) = 1$. If V is isotropic, then $\#(l(A) \bmod \Omega(V)) = 1$ if either A is anisotropic and $\langle A\rangle^*$ is isotropic or A is non-zero isotropic,* $\dim V \geqq 3$ *and the Witt index of $V > 1$; 2 if either $\langle A\rangle^*$ is anisotropic or A is non-zero isotropic,* $\dim V \geqq 3$ *and the Witt index of $V = 1$; 4 if A is non-zero isotropic and* $\dim V = 2$.

Proof. Immediate by Theorem 2, Lemma 4 and the index relation $(K^* : K^{*2}) = 2$ for K the field of real numbers.

COROLLARY 5. *$\Omega(V) = O'(V)$ is length transitive if and only if the Witt index of $V \neq 1$.*

6. K is a local field. By a *local field* we mean a field K complete with respect to a discrete valuation having finite residue class field K'. Following O'Meara (see [6], § 63) we call K *dyadic* (*non-dyadic*) if the characteristic of K' is 2 (not 2). The values of two group indices germane to the issue are given in

LEMMA 5. *Let K be a local field (of characteristic $\neq 2$) and E/K a quadratic extension. Then $(K^* : N_{E/K}(E^*)) = 2$ and $(N_{E/K}(E^*) : K^{*2}) = 2q^{\operatorname{ord} 2}$ where q is the number of elements in K'.*

Proof. These are well known results of local class field theory. (See [2], Chapters 8 and 10 or [6], § 63.)

LEMMA 6. *Let V be a non-singular quadratic space over a local field K. If* $\dim V \geqq 3$ *then $\theta(O^+(V)) = K^*$. If K is dyadic, then $\Omega(V) = O'(V)$. If K is non-dyadic, then $\Omega(V) = O'(V)$ unless* $\dim V = 4$ *and V is anisotropic in which case $(O'(V) : \Omega(V)) = 2$.*

Proof. If $\dim V \geqq 4$ then it is well known that V assumes every possible length (see [8], Satz 15 and Satz 16) and the result follows. It is also obvious if V is 3-dimensional isotropic. If V is 3-dimensional anisotropic then there is exactly one class α of $K^* \bmod K^{*2}$ which is not achieved as a length. (See [8], Satz 14 and Satz 15.) If β is a class distinct from α and 1 (β exists since $(K^*: K^{*2}) \geqq 4$) then $\alpha\beta \neq \alpha$ and $\alpha = \beta \cdot \alpha\beta$. Since β and $\alpha\beta$ are classes assumed by V as lengths, $\alpha \in \theta(O^+(V))$ also. The remaining assertions constitute Theorem C of [7].

THEOREM 5. *Let V be a non-singular quadratic space of dimension $\geqq 2$ over a local field. Then $\#(l(A) \bmod O'(V)) = 1$ if either $\dim V \geqq 5$ or $\dim V = 4$ and A is anisotropic or $\dim V = 4$, A is non-zero isotropic and Witt index of $V = 2$; 2 if either $\dim V = 4$, A non-zero isotropic and Witt index of $V = 1$ or $\dim V = 3$ and A is anisotropic; $2q^{\mathrm{ord}\,2}$ if $\dim V = 2$ and V is anisotropic; $4q^{\mathrm{ord}\,2}$ if either $\dim V = 3$ and A is non-zero isotropic or $\dim V = 2$, V isotropic and A is anisotropic; $8q^{\mathrm{ord}\,2}$ if $\dim V = 2$ and A is non-zero isotropic.*

Proof. Immediate by Theorem 2, Corollary 3, Lemma 5 and Lemma 6.

COROLLARY 6. *$O'(V)$ is length-transitive if and only if $\dim V \geqq 5$ or $\dim V = 4$ and the Witt index of $V \neq 1$.*

Remark 2. For local fields the condition Witt index $\neq 1$ for 4-dimensional non-singular quadratic spaces is equivalent to the condition $d(V) \in K^{*2}$.

THEOREM 6. *Let V be a non-singular quadratic space of dimension $\geqq 2$ over a local field K. Then $\#(l(A) \bmod \Omega(V)) = \#(l(A) \bmod O'(V))$ unless K is non-dyadic, $\dim V = 4$ and V anisotropic in which case*

$$\#(l(A) \bmod \Omega(V)) = 2[\#(l(A) \bmod O'(V))].$$

Proof. Immediate from Theorem 2, Theorem 5 and Lemma 6.

COROLLARY 7. *$\Omega(V)$ is length-transitive if and only if $\dim V \geqq 5$ or $\dim V = 4$ and the Witt index of $V \neq 1$ if K is dyadic; the Witt index of $V = 2$ if K is non-dyadic.*

7. K is a global field. By a *global field* we mean either an algebraic number field or an algebraic function field in one variable with finite constant field. If $\mathfrak{p}$ is a prime of K we will denote the completion of K at $\mathfrak{p}$ by $K_{\mathfrak{p}}$

and if U is any subspace of V we will denote by $U_{\mathfrak{p}}$ the scalar extension of U by $K_{\mathfrak{p}}$. Of course $U_{\mathfrak{p}}$ becomes a quadratic space over $K_{\mathfrak{p}}$ in a natural manner. If R denotes any (possibly empty) set of real infinite primes of K then by $K^*(R)$ we denote the set of all elements of K^* which are positive at each $\mathfrak{p}$ of R.

LEMMA 7. *Let K be a global field (of characteristic $\neq 2$). Let S, T be sets of real infinite primes of K satisfying $S \subseteq T$. Let $a \notin K^{*2}$ and suppose $-a \in K^*(S)$ so that $N_{E/K}(E^*) \subseteq K^*(S)$ where $E = K(a^{\frac{1}{2}})$. Then $(K^*(S) : K^*(T)) = 2^{\#(T-S)}$ and both $(K^*(S) : N_{E/K}(E^*))$ and $(N_{E/K}(E^*) : K^{*2})$ are infinite.*

Proof. If T is empty the first assertion is trivial. So suppose T is non-empty and let $T - S = \{\mathfrak{p}_1, \cdots, \mathfrak{p}_r\}$. For $i = 1, \cdots, r$ let f_i be the injection map of K into its completion $K_{\mathfrak{p}_i}$ at $\mathfrak{p}_i$. Define a map

$$f\colon K^*(S) \to (K_{\mathfrak{p}_1}{}^*/K_{\mathfrak{p}_1}{}^{*2}) \times \cdots \times (K_{\mathfrak{p}_r}{}^*/K_{\mathfrak{p}_r}{}^{*2})$$

by $f(x) = (f_1(x)K_{\mathfrak{p}_1}{}^{*2}, \cdots, f_r(x)K_{\mathfrak{p}_r}{}^{*2})$. f is obviously a morphism. By the weak approximation theorem (see [2], Theorem 8 of Chapter 1) it easily follows that f is surjective. Clearly $\operatorname{Ker} f = K^*(T)$ and the first assertion is proved. Suppose $(K^*(S) : N_{E/K}(E^*))$ was finite. Let $b_1, \cdots, b_n$ be a complete set of representatives for $K^*(S)/N_{E/K}(E^*)$ in $K^*(S)$ and suppose b_1 represents $N_{E/K}(E^*)$. For each $i > 1$ there exists a prime $\mathfrak{p}_i$ such that $b_i \notin N_{E_{\mathfrak{p}_i}/K_{\mathfrak{p}_i}}(E_{\mathfrak{p}_i}{}^*)$ where $E_{\mathfrak{p}_i} = K_{\mathfrak{p}_i}(a^{\frac{1}{2}})$. Note that $\mathfrak{p}_i \notin S$. Let $\mathfrak{q}$ be a prime such that $\mathfrak{q} \notin \{\mathfrak{p}_2, \cdots, \mathfrak{p}_n\} \cup S$ and also $a \notin K_{\mathfrak{q}}{}^{*2}$. Let $c_{\mathfrak{q}} \in K_{\mathfrak{q}}{}^*$ but $c_{\mathfrak{q}} \notin N_{E_{\mathfrak{q}}/K_{\mathfrak{q}}}(E_{\mathfrak{q}}{}^*)$. By the weak approximation theorem there exists $b \in K^*$ such that $b \equiv c_{\mathfrak{q}} \bmod K_{\mathfrak{q}}{}^{*2}$ and $b \equiv 1 \bmod K_{\mathfrak{p}}{}^{*2}$ for $\mathfrak{p} \in \{\mathfrak{p}_2, \cdots, \mathfrak{p}_n\} \cup S$. Thus $b \in K^*(S)$, $b \notin N_{E/K}(E^*)$ and $bb_i \equiv b_i \bmod N_{E_{\mathfrak{p}_i}/K_{\mathfrak{p}_i}}(E_{\mathfrak{p}_i}{}^*)$ for $i = 2, \cdots, n$. Thus $bb_i \notin N_{E/K}(E^*)$ for $i = 1, \cdots, n$ which contradicts the choice of $b_1, \cdots, b_n$ as a complete set of representatives.

Finally suppose $(N_{E/K}(E^*) : K^{*2})$ is finite. Let $c_1, \cdots, c_n$ be a complete system of representatives for $N_{E/K}(E^*)/K^{*2}$ in $N_{E/K}(E^*)$ and suppose c_1 represents K^{*2}. For each $i > 1$ there exists a finite prime $\mathfrak{p}_i$ such that $c_i \notin K_{\mathfrak{p}_i}{}^{*2}$. Let $\mathfrak{q}$ be a finite prime such that $\mathfrak{q} \neq \mathfrak{p}_i$ for $i = 2, \cdots, n$. Select $b_{\mathfrak{q}} \in N_{E_{\mathfrak{q}}/K_{\mathfrak{q}}}(E_{\mathfrak{q}}{}^*)$ but $b_{\mathfrak{q}} \notin K_{\mathfrak{q}}{}^{*2}$. There exist $x_{\mathfrak{q}}, y_{\mathfrak{q}} \in K_{\mathfrak{q}}$ such that $x_{\mathfrak{q}}{}^2 - ay_{\mathfrak{q}}{}^2 = b_{\mathfrak{q}}$. By the weak approximation theorem we can find $x, y \in K$ such that x is very close to $x_{\mathfrak{q}}$ in $K_{\mathfrak{q}}$ and very close to 1 in $K_{\mathfrak{p}_i}$ for $i = 2, \cdots, n$ and y is very close to $y_{\mathfrak{q}}$ in $K_{\mathfrak{q}}$ and very close to 0 in $K_{\mathfrak{p}_i}$ for

$i=2,\cdots,n$. Set $c=x^2-ay^2$. Then $c\in N_{E/K}(E^*)$. By the continuity of addition and multiplication in $K_{\mathfrak{q}}$, $c\equiv b_{\mathfrak{q}} \bmod K_{\mathfrak{q}}^{*2}$ if our approximations are good enough. Hence $c\notin K^{*2}$. And $c\in K_{\mathfrak{p}_i}^{*2}$ for $i=2,\cdots,n$ if our approximations are sufficiently good. But then $cc_i\equiv c_i \bmod K_{\mathfrak{p}_i}^{*2}$ whence $cc_i\notin K^{*2}$ for $i=1,\cdots,n$. This contradicts the fact that $c_1,\cdots,c_n$ constitute a complete set of representatives and the Lemma is proved.

LEMMA 8. *Let V be a non-singular quadratic space over a global field. If $\dim V\geqq 3$, then $\theta(O^+(V))=K^*(S)$ where S consists of those real infinite primes $\mathfrak{p}$ for which $V_{\mathfrak{p}}$ is anisotropic. If $\dim V\neq 4$, then $\Omega(V)=O'(V)$. If $\dim V=4$ then $O'(V)/\Omega(V)\cong\prod_{\mathfrak{p}\in R}O'(V_{\mathfrak{p}})/\Omega(V_{\mathfrak{p}})$ where R is the set of finite non-dyadic $\mathfrak{p}$ for which $V_{\mathfrak{p}}$ is anisotropic.*

Proof. These assertions are due to Kneser. (See [5] Satz A and Satz B for proofs of the first two assertions and [7] Theorem B for a proof of the last assertion.) Although they are stated only for algebraic number fields, the argument also works in the (easier) function field case.

THEOREM 7. *Let V be a non-singular quadratic space of dimension $\geqq 2$ over a global field. Let S denote the set of all real infinite primes $\mathfrak{p}$ for which $V_{\mathfrak{p}}$ is anisotropic. If A is an anisotropic vector, let $T(A)$ denote the set of all real infinite primes $\mathfrak{p}$ for which $\langle A\rangle^*_{\mathfrak{p}}$ is anisotropic. Then $\#(l(A) \bmod O'(V))=2^{\#(T(A)-S)}$ if either $\dim V\geqq 4$ or $\dim V=3$ and $\langle A\rangle^*$ is isotropic; ∞ otherwise. If A is a non-zero isotropic vector and $\dim V\geqq 3$, let $T(A)$ denote the set of all real infinite primes $\mathfrak{p}$ for which $W_{\mathfrak{p}}$ is anisotropic where $\langle A\rangle^*=\langle A\rangle\perp W$, W non-singular. Then $\#(l(A) \bmod O'(V))=2^{\#(T(A)-S)}$ if either $\dim V\geqq 5$ or $\dim V=4$ and the Witt index of $V=2$; ∞ otherwise.*

Proof. A straightforward application of Theorem 2, Corollary 3, Lemma 7 and Lemma 8.

COROLLARY 8. *$O'(V)$ is length-transitive if and only if the Witt index of $V_{\mathfrak{p}}\neq 1$ for all real infinite $\mathfrak{p}$ and either $\dim V\geqq 5$ or $\dim V=4$ and the Witt index of $V\neq 1$.*

THEOREM 8. *Let V be a non-singular quadratic space of dimension $\geqq 2$ over a global field. Then $\#(l(A) \bmod \Omega(V))=\#(l(A) \bmod O'(V))$ if $\dim V\neq 4$. If $\dim V=4$, then $\#(l(A) \bmod \Omega(V))=2^r[\#(l(A) \bmod O'(V))]$ where r is the number of finite non-dyadic $\mathfrak{p}$ for which $V_{\mathfrak{p}}$ is anisotropic.*

Proof. Immediate from Theorem 2, Theorem 7 and Lemma 8.

COROLLARY 9. *$\Omega(V)$ is length-transitive if and only if the Witt index of $V_{\mathfrak{p}} \neq 1$ for all real infinite $\mathfrak{p}$ and either* $\dim V \geqq 5$ *or* $\dim V = 4$, *the Witt index of $V \neq 1$ and $V_{\mathfrak{p}}$ is isotropic at all finite non-dyadic $\mathfrak{p}$.*

INSTITUTE FOR DEFENSE ANALYSES,
SYRACUSE UNIVERSITY.

REFERENCES.

[1] E. Artin, *Geometric Algebra*, New York, 1957.

[2] ———, *Algebraic numbers and algebraic functions I*, Princeton University, New York University, 1950-1951.

[3] J. Dieudonné, *Sur les groupes classiques*, Paris, 1948.

[4] ———, *La géométrie des groupes classiques*, Berlin, 1955.

[5] M. Kneser, "Orthogonale Gruppen über algebraischen Zahlkörpern," *Journal für die reine und angewandte Mathematik*, vol. 196 (1956), pp. 213-220.

[6] O. T. O'Meara, *Introduction to Quadratic Forms*, Berlin, 1963.

[7] B. Pollak, "4-dimensional orthogonal groups over algebraic number fields," *Journal für die reine und angewandte Mathematik*, vol. 211 (1962), pp. 176-178.

[8] E. Witt, "Theorie der quadratischen Formen in beliebigen Körpern," *Journal für die reine und angewandte Mathematik*, vol. 176 (1937), pp. 31-44.

CONTRACTIBLE SEMIGROUPS HAVE NO SELF-LINKED COMPACT SUBGROUPS.*

By PAUL S. MOSTERT.[1]

Wallace has asked the following question [2, p. 324]: If Euclidean n-space R^n is supplied with a continuous associative multiplication with identity, and G is a compact connected subgroup of R^n under the given multiplication, and contains the identity, can G be self-linked?[2] M. L. Curtis [1] has solved the question in the case where $n = 3$, the assumption that H contains the identity being dropped. Actually, the dimension and properties of Euclidean space (except the contractibility) have nothing to do with the question, nor does the existence of an identity as we see from the following proof of the title.

Proof. Let T be a contractible semigroup[3] and G a compact subgroup thereof. We denote the identity of G by e. Then $S = eTe$ is again a contractible semigroup and has e as its identity. Let $\pi\colon S \times I \rightarrow S$ be a homotopy, where I denotes the unit interval $[0,1]$, and π satisfies $\pi(x,0) = x$, $\pi(x,1) = x_0$ for every $x \in S$, x_0 being some fixed element of S. We have the following cases:

(a) Suppose S is compact. If S is a group, since it is contractible, it is a point. Otherwise, it contains a proper minimal ideal K [3]. Choose an element $k \in K$. Then $kS \cap G \subset K \cap G = \emptyset$.

If S is not compact, since $G\pi(G \times I)G$ is compact, there is an element $y \notin G\pi(G \times I)G$. Then we have either

(b) $yG\pi(G \times I)G \cap G = \emptyset$

or

(c) $yG\pi(G \times I)G \cap G \neq \emptyset$ in which case there is an element $z \in S$ such that $yz = e$. Then if $zG\pi(G \times I) \cap G \neq \emptyset$, by multiplying on the left by

* Received February 12, 1962.

[1] This work was supported by the National Science Foundation. The author wishes to thank A. M. Gleason for a pair of valuable comments concerning the proof.

[2] A set F is ***self-linked*** in X if there is no homotopy δ_t of F to a point such that $\delta_t(F) \cap F \neq \emptyset$ if and only if $t = 0$, where X is a contractible space.

[3] A ***semigroup*** is a Hausdorff space with a continuous, associative multiplication.

y, we have that $G\pi(G \times I)G$ contains yG contradicting the choice of y. Hence $zG\pi(G \times I)G \cap G = \emptyset$.

Now let p denote the point k, y, or z according as we have case (a), (b), or (c). Then we have

$$(*) \qquad pG\pi(G \times I)G \cap G = \emptyset .$$

There is an arc $a: [0,1] \to S$ with $a(0) = p$ and $a(1) = e$, since S is arcwise connected. Let t_0 be the smallest number such that $a(t_0) \in G$. Then t_0 is greater than zero. Say $a(t_0) = g_0$. Define now ϕ: $G \times I \to S$ by

$$\phi(h,t) = \begin{cases} a(t_0 - 2t_0 t) g_0^{-1} h, & 0 \leqq t \leqq \frac{1}{2} \\ p g_0^{-1} \pi(h, 2t-1), & \frac{1}{2} < t \leqq 1. \end{cases}$$

Clearly ϕ is continuous. Moreover $\phi(h,0) = h$, $\phi(h,1) = pg_0^{-1}x_0$.

Now suppose $\phi(h,t) \in G$ for some t, h. If $t \leqq \frac{1}{2}$, we have $a(t_0 - 2t_0 t) \in G$ so that $t = 0$. If $t > \frac{1}{2}$, we have $pg_0^{-1}\pi(h, 2t-1) \in G$ contradicting (*). Hence ϕ is a homotopy contracting G to a point in S, and hence in T, such that $\phi(h,t) \in G$ if and only if $t = 0$.

TULANE UNIVERSITY.

REFERENCES.

[1] M. L. Curtis, "Self-linked subgroups of semigroups," *American Journal of Mathematics*, vol. 81 (1959), pp. 889-892.

[2] A. D. Wallace, "Problems on semigroups," *Collectanea Mathematica* (Barcelona), vol. 8 (1961), pp. 223-224.

[3] ———, "The structure of topological semigroups," *Bulletin of the American Mathematical Society*, vol. 61 (1955), pp. 95-112.

ON A CUBIC EXPONENTIAL SUM IN THREE VARIABLES.*

By L. J. MORDELL.[1]

Let p be a prime and let $f(x,y,z)$ be a cubic polynomial with integer coefficients, which is not a function of less than three independent variables. Write $e(t) = e(2\pi it/p)$ and

$$S = \sum_{x,y,z=0}^{p-1} e(f(x,y,z)). \tag{1}$$

Very little is known about the exact order of magnitude of such sums for large p. An obvious conjecture for *general* $f(x,y,z)$ is

$$S = O(p^{\frac{3}{2}}), \tag{2}$$

where the constant implied in O is independent of the coefficients of $f(x,y,z)$. The result (2) is now proved for the special case when

$$f(x,y,z) = ax^3 + by^3 + cz^3 + dxyz,\ abcd \not\equiv 0, \tag{3}$$

except when $d^3 \equiv -27abc$ when we have an estimate $O(p^2)$.

The method of proof is similar [1] to that used in proving that

$$\sum_{x,y=0}^{p-1} e(ax^3 + by^3 + cxy) = O(p).$$

On multiplying (1) by $e(-aw^2)$ and summing for w, we have

$$\begin{aligned} \sum_{w=0}^{p-1} e(-aw^2)S &= \sum_{x,y,z,w=0}^{p-1} e(-aw^2 + ax^3 + by^3 + cz^3 + dxyz), \\ \left(\frac{-a}{p}\right) i^{((p-1)/2)^2}\sqrt{p}S &= \sum_{n,z=0}^{p-1} f(n)e(-an + cz^3) = S' \end{aligned} \tag{4}$$

say, where $f(n)$ is the number of solutions in x, y, w of

$$aw^2 \equiv ax^3 + by^3 + an + dxyz, \tag{5}$$

where all congruences are taken $(\operatorname{mod} p)$. It suffices now to prove that $S' = O(p^2)$.

* Received September 5, 1962.

[1] This work has been supported in part by the National Science Foundation, Washington, D. C.

In a recent paper, [2] I have found by elementary means simple closed expressions for the number of solutions of

$$w^2 \equiv x^3 + By^3 + CB^2 + DBxy, B \not\equiv 0. \tag{6}$$

Now

$$B \equiv b/a,\ CB^2 \equiv n,\ DB \equiv dz/a,$$

and so $D \equiv dz/b$, $C \equiv na^2/b^2$ with the usual meaning of $1/a$ mod p etc.

When $C \equiv 0$, i.e. $n \equiv 0$, $f(0) = p^2 + O(p)$. Consider first the cases when $p \equiv -1 \bmod 3)$, and $p \equiv 1 \pmod 3$ with B a cubic residue of p. From equation (16) of [2] we have when $n \not\equiv 0$,

$$f(n) = p^2 + p\left(\frac{3}{p}\right)\left(\frac{D^3 + 27C}{p}\right) + p\sum_r \left(\frac{6r - D}{p}\right), 2r^3 + Dr^2 + C \equiv \text{(} \tag{7}$$

where the brackets denote Jacobi-Legendre symbols. Hence

$$S' = \sum_{z=0}^{p-1} (p^2 + O(p))\,e(cz^3) + T,$$

$$T = \sum_{n=1, z=0}^{p-1} \left[p^2 + p\left(\frac{3}{p}\right)\left(\frac{d^3z^3/b^3 + 27na^2/b^2}{p}\right) + p\sum_r \left(\frac{6r - dz/b}{p}\right)\right] e(-an +$$

where $2r^3 + dzr/b + na^2/b^2 \equiv 0$.

The p^2 terms contribute 0 to S'. The term $O(p)\sum_{z=0}^{p-1} e(cz^3)$ gives $O(p^{3/2})$.

For the second term in T, put $m \equiv d^3z^3/b^3 + 27na^2/b^2$. It becomes

$$p\left(\frac{3}{p}\right)\sum_{m,z}{}' \left(\frac{m}{p}\right) e\left[\frac{-b^2m}{27a} + \left(\frac{d^3}{27ab} + c\right)z^3\right].$$

In the summation, since $n \not\equiv 0$, we must exclude $m \equiv d^3z^3/b^3$.

The sum in m is a Gaussian sum and gives $O(\sqrt{p})$. The sum in z gives $O(\sqrt{p})$ unless $d^3 + 27abc \equiv 0$ when it gives $O(p)$. The third sum in T is also $O(p)$ as we prove now. With a slight change in the variables, it suffices to show that

$$\sum_{U,V=0}^{p-1} \left(\frac{U}{p}\right) e(PU^3 + QU^2V + RUV^2 + SV^3) = O(p). \tag{8}$$

Put $V = UW$, then (8) becomes

$$\sum_{U,W=0}^{p-1} \left(\frac{U}{p}\right) e(U^3(P + QW + RW^2 + SW^3)).$$

On replacing U by λU^2, it suffices to show that

$$(9) \qquad S'' = \sum_{U,W=0}^{p-1} e(U^6(P+QW+RW^2+SW^3)) = O(p).$$

Let

$$\sum_{U=0}^{p-1} e(r_1 U^6) = \mathcal{S}_1\sqrt{p}, \cdots, \sum_{U=0}^{p-1} e(r_6 U^6) = \mathcal{S}_6\sqrt{p},$$

where $\mathcal{S}_1 = O(1), \cdots, \mathcal{S}_6 = O(1)$, and $r_1, \cdots, r_6$ are representatives of the sixth power residue classes of p. Let M_1 etc. denote the number of solutions of $r_1 Z^6 \equiv P + QW + RW^2 + SW^3$. Then $M_1 = p + O(\sqrt{p})$ by Weil's theorem since from (2) the cubic in (9) cannot be a multiple of a cube of a linear form. Hence

$$6S'' = \sqrt{p}\mathcal{S}_1(p + O(\sqrt{p}) + \cdots + \sqrt{p}\mathcal{S}_6(p + O(\sqrt{p})) = O(p).$$

since $\mathcal{S}_1 + \cdots + \mathcal{S}_6 = 0$.

Suppose next that $p \equiv 1 \pmod 3$ and that B is a non-cubic residue of p. Now from equation (18) of [2] when $n \not\equiv 0$,

$$(10) \qquad f(n) = p^2 - \tfrac{3}{2}\left(\frac{C}{p}\right) p + 3\left(\frac{C}{p}\right) pN_2 - \tfrac{1}{2}p\left(\frac{3}{p}\right)\left(\frac{D^3+27C}{p}\right) - \tfrac{1}{2}p\sum_r \left(\frac{6r-D}{p}\right)$$

taken over $2r^3 + Dr^2 + C \equiv 0$ with N_2 solutions.

The first and the last two terms in (8) have already been dealt with giving a contribution $O(p^2)$ unless $d^3 + 27abc \equiv 0$ when they give $O(p^{\frac{5}{2}})$. The second term contributes

$$-\tfrac{3}{2}p\sum_{n=1,z=0}^{p-1}\left(\frac{n}{p}\right) e(-an + cz^3) = O(p^2),$$

The third term gives

$$3p\sum_{n,z=0}^{p-1}\left(\frac{n}{p}\right) N_2 e(-an + cz^3)$$

where

$$2r^3 + \frac{dz}{b} r^2 + \frac{na^2}{b^2} \equiv 0$$

This is $O(p^2)$, for on substituting for n, we are led to a series such as (8). This finishes the proof.

The same method shows that if

$$(11) \qquad S_1 = \sum_{x,y,z} \chi(z) e(ax^3 + by^3 + cz^3 + dxyz),$$

where $\chi(z)$ is a multiplicative non principal character mod p, then also $S_1 = O(p^{\frac{3}{2}})$. Now the z sum arising from (7) is

$$\sum_z \chi(z) e\left((c + \frac{d^3}{27ab}) z^3 \right).$$

The sum in z is zero if $d^3 \equiv -27abc$, and also when $d^3 \not\equiv -27abc$ when $p \equiv -1 \pmod 3$. If $p \equiv 1 \pmod 3$, the sum is $O(\sqrt{p})$.

On noting how (11) arises, the linear term there is not a factor of the cubic term. Now (8) becomes

$$\sum_{U,V=0}^{p-1} \left(\frac{U}{p}\right) \chi(V) e(PU^3 + QU^2V + RUV^2 + SV^3),$$

and with $V \equiv UW$, this is

$$\sum_{U,W=0}^{p-1} \left(\frac{U}{P}\right) \chi(U)\chi(W) e(U^3(P + QW + RW^2 + SW^3))$$

It is easily seen that this sum is also $O(p)$ since the sum in U is $\chi'(P + QW + RW^2 + SW^3)$ where χ' is a multiplicative character.

UNIVERSITY OF ARIZONA, TUCSON, ARIZONA
ST. JOHN'S COLLEGE, CAMBRIDGE, ENGLAND.

REFERENCES.

[1] L. J. Mordell, "On a cubic exponential sum in two variables," to appear in the *Journal of the London Mathematical Society*.

[2] ———, "Cubic congruences in three variables III," to appear in the *Journal of the London Mathematical Society*.

NOTE ON THE REDUCTION OF INDUCED REPRESENTATIONS.*

By PATRICIA A. TUCKER.

1. **Introduction.** Let G be a finite group with a normal subgroup H. If L is an irreducible left $K(H)$-module, then one is interested in determining the components of the induced module $L^G = K(G) \otimes_{K(H)} L$. In this paper the indecomposable components of L^G are determined when K is an algebraically closed field of characteristic $p \geqq 0$. This is an extension of the author's earlier results [1] in which G was assumed to be a split extension of H by G/H. The development presented here will depend heavily upon the results in [1].

2. **Construction of the indecomposable components.** Let G be a finite group with normal subgroup H. Let G be an extension of H by B where the elements of B are $b_1 = 1, b_2, \cdots, b_n$. To each element of B, there corresponds a coset of H in G. Let the left coset representatives of H in G be $\bar{b}_1 = 1, \bar{b}_2, \cdots, \bar{b}_n$. If G is actually a split extension of H by G/H, consider B as a subgroup of G and take $\bar{b}_i = b_i$, $i = 1, \cdots, n$. Thus,

$$G = H + \bar{b}_2 H + \cdots + \bar{b}_n H.$$

Each $b \in B$ defines an automorphism of H by $h \to (\bar{b})^{-1} h \bar{b} = h^b$, for every $h \in H$.

Multiplication of elements of G is given by

$$\bar{b}h \cdot \bar{b}'h' = \overline{bb'}(b, b') h^{b'} h', \text{ where } b, b' \in B \text{ and } h, h' \in H. \tag{1}$$

$\{(b_i, b_j) \mid 1 \leqq i, j \leqq n\}$ is the factor set of the extension. Since $b_1 = 1$, $(1, b_j) = 1 = (b_j, 1)$, for every j. From the associative law for multiplication of elements of G, it follows that for every i, j, and k

$$(b_i b_j, b_k)(b_i, b_j)^{b_k} = (b_i, b_j b_k)(b_j, b_k). \tag{2}$$

Let T be an irreducible representation of H over an algebraically closed field K of characteristic $p \geqq 0$. Let L be a left $K(H)$-module which is a representation module for T.

The conjugate representation of T under $x \in G$ is denoted by $T^{(x)}$ and is such that $T^{(x)}(h) = T(x^{-1}hx)$, for all $h \in H$.

* Received September 15, 1962.

LEMMA 1. *Let* $S = \{b \in B \mid T^{(\bar{b})}$ *is equivalent to* $T\}$. *Then,* S *is a subgroup of* B.

If G *is a split extension of* H *by* G/H, *then* $\bar{b} = b$ *and* S *as defined here is the same* S *as in* [1].

Proof. If $b \in S$, then there exists a non-singular linear transformation D_b of L such that $D_b T^{(\bar{b})}(h) D_b^{-1} = T(h)$, for every $h \in H$, i.e., $D_b T^{(\bar{b})} D_b^{-1} = T$. The proof of the lemma follows immediately from the following:

$$(3) \qquad (T((b,b^{-1}))^{-1}D_b)^{-1}T^{(\overline{b^{-1}})}T((b,b^{-1}))^{-1}D_b = T$$

$$(4) \qquad T^{(\overline{bb'})} = T((b,b'))D_{b'}^{-1}D_b^{-1}TD_bD_{b'}T((b,b'))^{-1}.$$

For every $b \in S$, select a linear transformation D_b of L such that $D_b T^{(\bar{b})} D_b^{-1} = T$ and $D_1 = 1_L$. Then, for every $l \in L$, $h \in H$, $D_b h^b D_b^{-1} l = hl$. Let b and $b' \in S$. From (4) it follows that

$$T = (D_bD_{b'}T((b,b'))^{-1})T^{(\overline{bb'})}(D_bD_{b'}T((b,b'))^{-1})^{-1}.$$

By definition of $D_{bb'}$, $T = D_{bb'}T^{(\overline{bb'})}D_{bb'}^{-1}$. Thus, by Schur's Lemma

$$(5) \qquad D_bD_{b'}T((b,b'))^{-1} = \beta(b,b')D_{bb'},$$

where $\beta(b,b') \in K^* = \{k \in K \mid k \neq 0\}$.

LEMMA 2. β *is a factor set of* S.

Proof. One only needs to prove that for every $a, b, c \in S$, $\beta(a,bc)\beta(b,c) = \beta(a,b)\beta(ab,c)$. Now,

$$\begin{aligned}
\beta(a,b)\beta(ab,c) &= D_aD_bT((a,b))^{-1}D_{ab}^{-1}D_{ab}D_cT((ab,c))^{-1}D_{abc}^{-1}\\
&= D_aD_bT((a,b))^{-1}D_cT((a,bc)(b,c)((a,b)^c)^{-1})^{-1}D_{abc}^{-1}\\
&= D_aD_bT((a,b))^{-1}D_cT((a,b)^c)T((b,c))^{-1}T((a,bc))^{-1}D_{abc}^{-1}\\
&= D_aD_bT((a,b))^{-1}T((a,b))D_cT((b,c))^{-1}D_{bc}^{-1}D_a^{-1}\beta(a,bc)\\
&= D_a\beta(b,c)D_{bc}D_{bc}^{-1}D_a^{-1}\beta(a,bc) = \beta(a,bc)\beta(b,c).
\end{aligned}$$

It can be proved that if any other selection is made for the D_b's, then the resulting factor set is equivalent to β. Conversely, let β' be a factor set of S which is equivalent to β. Then, there exists a function $\rho: S \to K^*$ such that $\beta'(b,a) = \beta(a,b)\rho(ab)\rho(a)^{-1}\rho(b)^{-1}$, for every $a, b \in S$. Let $C_b = \rho(b)D_b$. Then, $C_bT^{(\bar{b})}C_b^{-1} = T$ and $C_bC_{b'}T((b,b'))^{-1} = \beta'(b,b')C_{bb'}$. Thus, T uniquely determines up to equivalence a factor set of S.

Let $(KS)_{\beta^{-1}}$ denote the β^{-1}-twisted group algebra of S, i.e., the crossed product of K and S with factor set β^{-1} (See [1].) Since β is equivalent to β' if and only if β^{-1} is equivalent to $(\beta')^{-1}$, it follows that T determines up to equivalence $(KS)_{\beta^{-1}}$. The elements of $(KS)_{\beta^{-1}}$ will be denoted by s, where

$$s = \sum_{b \in S} \xi_s(b)b.$$

Let a_γ, $\gamma=1,\cdots,r=[B:S]$ and $a_1=1$, be a complete set of left coset representatives of S in B. Then, $B=S+a_2S+\cdots+a_rS$. It can be proved that $T^{(\bar{a}_\gamma)}$ is inequivalent to $T^{(\bar{a}_\lambda)}$, for $\gamma\neq\lambda$.

The left $K(G)$-module L^G for T^G has dimension $[B:1](L:K)$ and is generated as a vector space over K by the elements

$$\{\overline{a_\gamma b}\otimes l \mid 1\leqq\gamma\leqq r,\ b\in S,\ l\in L\}.$$

For every $\gamma=1,\cdots,r$, $s\in(KS)_{\beta^{-1}}$, and $l\in L$, define an element $c(a_\gamma,s,l)\in L^G$ by

$$c(a_\gamma,s,l)=\sum_{b\in S}\xi_s(b)\overline{a_\gamma b}\otimes T((a_\gamma,b))D_b{}^{-1}l.$$

If G is a split extension of H by G/H, then $(a_\gamma,b)=1$ and $c(a_\gamma,s,l)$ is the same as $c(a_\gamma,s,l)$ in [1]. The $c(a_\gamma,s,l)$'s satisfy the following properties:

Property 1. $c(a_\gamma,s+s',l)=c(a_\gamma,s,l)+c(a_\gamma,s',l)$.

Property 2. $c(a_\gamma,s,l+l')=c(a_\gamma,s,l)+c(a_\gamma,s,l')$.

Property 3. If $k\in K$, then

$$k(c(a_\gamma,s,l))=c(a_\gamma,ks,l)=c(a_\gamma,s,kl).$$

Property 4. If $h\in H$, then $h(c(a_\gamma,s,l))=c(a_\gamma,s,h^{a_\gamma}l)$.

Property 5. If $b\in S$, then

$$\bar{b}(c(a_\gamma,s,l))=c(a_\lambda,b'\cdot s,D_{b'}T((a_\lambda,b'))^{-1}T((b,a_\gamma))l),$$

where $ba_\gamma=a_\lambda b'$ and $b'\cdot s=\sum_{b\in S}\xi_s(b)\beta(b',b)^{-1}b'b$, i.e., the product of b' and s in $(KS)_{\beta^{-1}}$.

Property 6. If a_β is a coset representative of S in B, then

$$\overline{a_\beta}(c(a_\gamma,s,l))=c(a_\lambda,b'\cdot s,D_{b'}T((a_\lambda,b'))^{-1}T((\mathrm{b}_1,a_\lambda))l),$$

where $a_\beta a_\gamma=a_\lambda b'$.

Proof of Property 4.

$$\begin{aligned}
h(c(a_\gamma,s,l)) &= h(\sum_{b\in S}\xi_s(b)\overline{a_\gamma b}\otimes T((a_\gamma,b))D_b{}^{-1}l)\\
&=\sum_{b\in S}\xi_s(b)h\overline{a_\gamma b}\otimes T(,a_\gamma,b))D_b{}^{-1}l\\
&=\sum_{b\in S}\xi_s(b)\overline{a_\gamma b}(1,a_\gamma b)h^{a_\gamma b}\otimes T((a_\gamma,b))D_b{}^{-1}l\\
&=\sum_{b\in S}\xi_s(b)\overline{a_\gamma b}\otimes h^{a_\gamma b}T((a_\gamma,b))D_b{}^{-1}l\\
&=\sum_{b\in S}\xi_s(b)\overline{a_\gamma b}\otimes T((a_\gamma,b))D_b{}^{-1}h^{a_\gamma}l=c(a_\gamma,s,h^{a_\gamma}l).
\end{aligned}$$

Proof of Property 5. For convenience, replace b by b_1.

$$\begin{aligned}\overline{b_1}(c(a_\gamma, s, l)) &= \overline{b_1}(\sum_{b\in S}\xi_s(b)\overline{a_\gamma b}\otimes T((a_\gamma, b))D_b^{-1}l)\\ &= \sum_{b\in S}\xi_s(b)\overline{b_1}\,\overline{a_\gamma b}\otimes T((a_\gamma, b))D_b^{-1}l\\ &= \sum_{b\in S}\xi_s(b)\overline{b_1 a_\gamma b}(b_1, a_\gamma b)\otimes T((a_\gamma, b))D_b^{-1}l\\ &= \sum_{b\in S}\xi_s(b)\overline{a_\lambda b' b}\otimes T((b_1, a_\gamma b))T((a_\gamma, b))D_b^{-1}l.\end{aligned}$$

Now,

$$\begin{aligned}T((b_1, a_\gamma b))T((a_\gamma, b))D_b^{-1} &= T((b_1, a_\gamma b)(a_\gamma, b))D_b^{-1}\\ &= T((b_1 a_\gamma, b)(b_1, a_\gamma)^b)D_b^{-1} = T((a_\lambda b', b)(b_1, a_\gamma)^b)D_b^{-1}\\ &= T((a_\lambda, b'b)(b', b)((a_\lambda, b')^b)^{-1}(b_1, a_\gamma)^b)D_b^{-1}\\ &= T((a_\lambda, b'b))T((b', b))T^{(\bar b)}((a_\lambda, b')^{-1}(b_1, a_\gamma))D_b^{-1}\\ &= T((a_\lambda, b'b))T((b', b))D_b^{-1}T((a_\lambda, b')^{-1}(b_1, a_\gamma))\\ &= T((a_\lambda, b'b))\beta(b', b)^{-1}D_{b'b}^{-1}D_{b'}T((a_\lambda, b')^{-1}(b_1, a_\gamma)).\end{aligned}$$

Thus,

$$\begin{aligned}\overline{b_1}(c(a_\gamma, s, l)) &= \sum_{b\in S}\xi_s(b)\beta(b', b)^{-1}\overline{a_\lambda b' b}\\ &\quad\otimes (T((a_\lambda, b'b))D_{b'b}^{-1}D_{b'}T((a_\lambda, b'))^{-1}T((b_1, a_\gamma))l).\end{aligned}$$

The proof of Property 6 is similar.

Let I be a left ideal of $(KS)_{\beta^{-1}}$. Define $C(I)$ to be the K-subspace of L^G generated by the elements

$$\{c(a_\gamma, s, l) \mid 1 \leqq \gamma \leqq r, s \in (KS)_{\beta^{-1}}, l \in L\}.$$

From Properties 4, 5, and 6, $C(I)$ is a left $K(G)$-module. It can be shown (as in [1]) that $(C(I):K) = [B:S](I:K)(L:K)$ and that if $\{s_k \mid 1 \leqq k \leqq (I:K)\}$ forms a basis for I and if $\{l_\mu \mid 1 \leqq \mu \leqq (L:K)\}$ forms a basis for L, then

$$\{c(a_\gamma, s_k, l_\mu) \mid 1 \leqq \gamma \leqq r, 1 \leqq k \leqq (I:K), 1 \leqq \mu \leqq (L:K)\}$$

forms a basis for $C(I)$.

THEOREM 1. *Let $I_1, \cdots, I_t$ be left ideals of $(KS)_{\beta^{-1}}$ such that $(KS)_{\beta^{-1}} = I_1 \oplus \cdots \oplus I_t$. Then, $L^G = C(I_1) \oplus \cdots \oplus C(I_t)$.*

The proof of this theorem is essentially identical to the proof of Theorem 4 in [1].

THEOREM 2. *Let I_1 and I_2 be left ideals of $(KS)_{\beta^{-1}}$. Then, there exists*

a K-isomorphism θ of $\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$ *onto* $\mathrm{Hom}_{K(G)}(C(I_1), C(I_2))$ *such that if* $P \in \mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$, *then* $P(c(a_\gamma, s, l)) = c(a_\gamma, Ps, l)$, *for every* $c(a_\gamma, s, l) \in C(I_1)$. *If* $I_1 = I_2$, *then θ is an algebra isomorphism.*

The proof of this theorem follows the same pattern as the proof of Theorem 5 in [1]. Only a few minor changes are necessary to extend that proof to this case.

The following corollaries may be proved as in [1].

COROLLARY 1. *I is indecomposable if and only if $C(I)$ is indecomposable.*

COROLLARY 2. *L^G is indecomposable if and only if $(KS)_{\beta^{-1}}$ is indecomposable.*

COROLLARY 3. *If the characteristic p of K does not divide $[G\colon 1]$, then I is irreducible if and only if $C(I)$ is irreducible.*

COROLLARY 4. *Suppose p does not divide $[G\colon 1]$. Let I_1 and I_2 be irreducible. Then, I_1 and I_2 are $(KS)_{\beta^{-1}}$-isomorphic if and only if $C(I_1)$ and $C(I_2)$ are $K(G)$-isomorphic.*

3. Further results. If $S \neq B$, then $C(I)$ is $K(G)$-isomorphic to a left $K(G)$-module which is induced from a left $K(\bar{S}H)$-module, where $\bar{S}H$ is a proper subgroup of G.

Let $\bar{S} = \{\bar{b} \mid b \in S\}$. Then, $\bar{S}H = \{\bar{b}h \mid b \in S, h \in H\}$ is a subgroup of G. $\{\overline{a_\gamma} \mid 1 \leqq \gamma \leqq r\}$ forms a complete set of left coset representatives of $\bar{S}H$ in G. Thus, $G = \bar{a}_1\bar{S}H + \cdots + \bar{a}_r\bar{S}H$.

Let I' be the K-subspace of L^G generated by $\{c(1, s, l) \mid s \in I, l \in L\}$. Then,

$$\bar{b}h(c(1, s, l)) = \bar{b}(c(1, s, hl)) = c(1, b \cdot s, D_b hl).$$

Thus, I' is a left $K(\bar{S}H)$-module. It is easily checked that the linear transformation ϕ of $(I')^G$ into $C(I)$ such that $\phi\colon \overline{a_\gamma} \otimes c(1, s, l) \to c(a_\gamma, s, l)$ is a $K(G)$-isomorphism of $(I')^G$ onto $C(I)$. Thus, $(I')^G$ and $C(I)$ are isomorphic as left $K(G)$-modules.

As in the case of a split extension (see [1, § 4]), it can be proved that I' is $K(\bar{S}H)$-isomorphic to the (inner) tensor product of two modules which afford projective representations of $\bar{S}H$. Theorem 7 in [1] then holds with the condition that G is the split extension of H by G/H removed and α replaced by β.

4. Summary.

THEOREM 3. *Let G be a finite group with normal subgroup H, G/H isomorphic to B. Let T be an irreducible representation of H over an algebraically closed field K of characteristic p. Let L be a representation module for T. Let $S = \{b \in B \mid T^{(\bar{b})}$ is equivalent to $T\}$. For each $b \in B$, define a non-singular linear transformation D_b of L such that $D_b T^{(\bar{b})} D_b^{-1} = T$ and $D_1 = 1_L$. Define $\beta(b, b') \in K^*$ by $D_b D_{b'} T((b,b'))^{-1} = \beta(b,b') D_{bb'}$. β is a factor set of S. Let $(KS)_{\beta^{-1}}$ denote the β^{-1}-twisted group algebra of S.*

Each left ideal I of $(KS)_{\beta^{-1}}$ determines a left $K(G)$-module $C(I)$ of L^G such that

1) $(C(I):K) = (I:K)[B:S](L:K)$.
2) There exists a left $K(SH)$-module I' such that I' is $K(SH)$-isomorphic to the tensor product of two modules which afford projective representations of SH and $(I')^G$ is $K(G)$-isomorphic to $C(I)$.
3) If $I_1, \cdots, I_t$ are left ideals of $(KS)_{\beta^{-1}}$ such that
$$(KS)_{\beta^{-1}} = I_1 \oplus \cdots \oplus I_t,$$
then $L^G = C(I_1) \oplus \cdots \oplus C(I_t)$.
4) If I_1 and I_2 are left ideals of $(KS)_{\beta^{-1}}$, then $\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$ is K-isomorphic to $\mathrm{Hom}_{K(G)}(C(I_1), C(I_2))$. (If $I_1 = I_2$, then have an algebra isomorphism.)

Let $\mathfrak{J}_I$ denote the representation of G afforded by $C(I)$. Then, if $S \neq B$, $\mathfrak{J}_I$ is an induced representation. If $(KS)_{\beta^{-1}} = I_1 \oplus \cdots \oplus I_t$, where the I_i are indecomposable left ideals, then $T^G = \mathfrak{J}_{I_1} \oplus \cdots \oplus \mathfrak{J}_{I_t}$ is a decomposition of T^G into indecomposable representations. If p does not divide $[G:1]$, then the $\mathfrak{J}_{I_i}$ are irreducible and $\mathfrak{J}_{I_i}$ and $\mathfrak{J}_{I_j}$ are equivalent if and only if I_i and I_j are $(KS)_{\beta^{-1}}$-isomorphic.

UNIVERSITY OF ILLINOIS, URBANA, ILLINOIS.

REFERENCES.

[1] Patricia A. Tucker, "On the reduction of induced representations of finite groups," *American Journal of Mathematics*, vol. 84 (1962), pp. 400-420.

ON UNBOUNDED TOEPLITZ MATRICES.*[1]

By PHILIP HARTMAN.

1. **Introduction.** The function spaces L^p below refer to functions of a real variable ϕ on the interval $0 \leqq \phi \leqq 2\pi$. (Null ϕ-sets will be ignored without comment.) Let $f(\phi) \in L^2$ and

$$(1.1) \qquad f(\phi) \sim \sum_{n=-\infty}^{+\infty} f_n e^{in\phi}$$

and

$$(1.2) \qquad \| f \| = \left(\sum_{n=-\infty}^{\infty} | f_n |^2 \right)^{\frac{1}{2}} \geqq 0.$$

Correspondingly, let l^2 be the space of sequences $f = (\cdots, f_{-1}, f_0, f_1, \cdots)$ with a (finite) norm (1.2). Let $f = (f^-, f^+)$, where $f^\pm$ are the respective sequences $f^+ = (f_0, f_1, \cdots)$, $f^- = (\cdots, f_{-2}, f_{-1})$ and let the corresponding functions $f^\pm(\phi)$ be

$$(1.3) \qquad f^+(\phi) \sim \sum_{n=0}^{\infty} f_n e^{in\phi}, \qquad f^-(\phi) \sim \sum_{n=1}^{\infty} f_{-n} e^{-in\phi}.$$

The subspace of functions $f^+(\phi)$ of L^2 will be denoted by L^{2+} and the corresponding space of sequences f^+ by l^{2+}. When convenient, l^{2+} will be considered a subspace of l^2 by associating each element $f^+ \in l^{2+}$ with the doubly infinite sequence $(0, f^+) \in l^2$. Similarly, L^2 and l^2 or L^{2+} and l^{2+} will be identified.

Let $L = L(f)$ and $T = T(f)$ denote the Laurent and Toeplitz matrices belonging to f, i. e., $L = (f_{n-m})$, where $n, m = 0, \pm 1, \cdots$, and $T = (f_{n-m})$, where $n, m = 0, 1, \cdots$. The symbols $L(f)$, $T(f)$ will also represent the corresponding operators of the spaces l^2, l^{2+} or L^2, L^{2+} into themselves with the largest domain that one can associate with the respective matrices.

If $f(\phi)$ is real-valued, i. e., $f_n = \bar{f}_{-n}$, then L and T are formally Hermitian matrices. L, in fact, is a self-adjoint operator. It was pointed out in [4], p. 879, however, that T need not be self-adjoint. A sufficient condition for T to be self-adjoint is that the real function $f(\phi)$ be bounded or half-bounded; [4].

The motivation for this paper was an investigation of conditions, necessary and/or sufficient, on a real-valued $f(\phi) \in L^2$ for $T(f)$ to be self-adjoint.

* Received October 29, 1962.

[1] This research was supported by the Air Force Office of Scientific Research.

Actually, only the last section deals directly with this problem and depends on applications of the earlier sections which involve $T(f)$ without the restriction that $f(\phi)$ is real-valued.

The paper depends, in part, on the answer to the following question: Let $f^+(\phi)$ in (1.3) be in L^{2+} and $M(f^+) = \text{c.l.m.}(f^+, f^+e^{i\phi}, \cdots)$, the subspace (= closed, linear manifold) spanned by $f^+e^{in\phi}$ for $n = 0, 1, \cdots$. When is $M(f^+) = L^{2+}$? This question was raised and investigated in Hartman [3]. Its complete answer was subsequently given in a well-known paper by Beurling [1]. He showed, among other things, that f^+ has an essentially unique factorization $f^+(\phi) = f_0^+(\phi)f_1^+(\phi)$ into an "inner" and "outer" function, that $|f_0^+(\phi)| = 1$, that $M(f^+) = M(f_0^+) = f_0^+(\phi) \cdot L^{2+}$, and that $M(f^+) = L^{2+}$ if and only if $f_0^+ \equiv \text{const}$. The function f_0^+ was given analytically and also geometrically in terms of projections in L^{2+}. Halmos paper [2] contains an elegant geometrical derivation of this result (but does not supply f_0^+ analytically). The arguments below depend both on Beurling's results and on Halmos's results on a more general problem.

By the use of these arguments, questions concerning the spectrum of $T(f)$ are reduced in part to an investigation of $T(f_0)$ for a suitable $f_0(\phi)$ with $|f_0(\phi)| = 1$. When $f(\phi)$, $1/f(\phi)$ are both bounded, such a reduction was made by Widom [11] with $f_0 = f/|f|$.

2. The subspaces $M(f)$. Let U denote the unitary operator of L^2 onto itself defined by $Uf(\phi) = f(\phi)e^{i\phi}$; so that $U^*f(\phi) = U^{-1}f(\phi) = f(\phi)e^{-i\phi}$. A subspace M of L^2 is an *invariant* subspace of U if $U(M) \subset M$ and *reduces* U if $U(M) \subset M$, $U^*(M) \subset M$. An invariant subspace $M \neq \{0\}$ of U is called *irreducible* if it contains no non-trivial ($\neq \{0\}$) subspace which reduces U.

LEMMA 2.1 (Halmos [2]). *An invariant subspace $M \neq \{0\}$ of U is irreducible if and only if there exists a function $f_0(\phi) \in L^2$ such that $|f_0(\phi)| = 1$,*

$$M = M(f_0) = f_0(\phi) \cdot L^{2+}. \tag{2.1}$$

The function $f_0(\phi)$ is unique up to a constant factor of absolute value 1.

If $f(\phi) \in L^2$, $M(f)$ is the subspace

$$M(f) = \text{c.l.m.}\,(f, fe^{i\phi}, fe^{2i\phi}, \cdots) \tag{2.2}$$

spanned by $f, fe^{i\phi}, \cdots$. In (2.1), $f_0(\phi) \cdot L^{2+}$ denotes the set of functions $f(\phi)$ of the form $f_0(\phi)x^+(\phi)$, where $x^+(\phi) \in L^{2+}$.

THEOREM 2.1. *Let* $0 \not\equiv f(\phi) \in L^2$. *Then* $M(f)$ *is irreducible if and only if* $f(\phi) \neq 0$ *and*

$$\log|f(\phi)| \in L^1; \tag{2.3}$$

this is the case if and only if $f(\phi)$ *has a factorization*

$$f(\phi) = f_0(\phi) f_1^+(\phi), \text{ where } |f_0(\phi)| = 1 \tag{2.4}$$

and $f_1^+(\phi)$ *is an outer function in the sense of Beurling. In the factorization* (2.4), $f_0(\phi)$ *is unique up to a constant factor of absolute value* 1. *Furthermore,* $M(f) = M(f_0)$.

Before beginning the proof, several definitions will be recalled. Let

$$\alpha(\phi) \sim \sum_{n=-\infty}^{+\infty} \alpha_n e^{in\phi} \tag{2.5}$$

be of class L^1. The conjugate function $\beta(\phi)$ of $\alpha(\phi)$ is the function defined by the Abel sum of the series

$$\beta(\phi) \sim -i\sum_{n=1}^{\infty} \alpha_n e^{in\phi} + i\sum_{n=1}^{\infty} \alpha_{-n} e^{-in\phi} \tag{2.6}$$

or, equivalently, by the principal value

$$\beta(\phi) = \frac{1}{2\pi} PV \int_0^{2\pi} \alpha(\theta) \cot \tfrac{1}{2}(\phi - \theta)\, d\theta;$$

cf. [12]. If $\alpha(\phi)$ is real-valued (i. e., $\alpha_n = \bar{\alpha}_{-n}$) and

$$A(z) = \alpha_0 + \frac{1}{2\pi} \int_0^{2\pi} \alpha(\phi)(e^{i\phi} + z)(e^{i\phi} - z)^{-1} d\phi, \tag{2.7}$$

then $A(z)$ is regular for $|z| < 1$ and $A(re^{i\phi}) \to \alpha(\phi) + i\beta(\phi)$ as $r \to 1$. A regular function $F(z)$ on $|z| < 1$ is called an outer function if there exists a real-valued function $\alpha(\phi) \in L^1$ and a real constant α such that $F(z) = \exp(A(z) + i\alpha)$. (In particular, an outer function does not vanish for $|z| < 1$.) Correspondingly, the boundary function $f^+(\phi) = F(e^{i\phi}) = \exp[\alpha(\phi) + i\beta(\phi) + i\alpha]$ will be called an outer function. A function $F_0(z)$ regular for $|z| < 1$, $|F_0(z)| \leqq 1$ and $|F_0(e^{i\phi})| = 1$ will be called an inner function. Correspondingly, the boundary function $f_0^+(\phi) = F_0(e^{i\phi})$ is called an inner function. (Thus $f_0^+(\phi)$ is an inner function if and only if $f_0^+(\phi) \in L^{2+}$ and $|f_0(\phi)| = 1$.)

Let $0 \not\equiv F(z) \in H^p$ for some $p > 0$, i. e., $F(z)$ is regular for $|z| < 1$ and $\int |F(re^{i\phi})|^p d\phi \leqq \text{const.} < \infty$ for $0 < r < 1$. Then $\alpha(\phi) = \log|F(e^{i\phi})| \in L^1$;

F. Riesz [6]. The outer function $F_1(z) = \exp(A(z) + i\alpha)$ satisfies $|F(z)| \leqq |F_1(z)|$ and $|F(e^{i\phi})| = |F_1(e^{i\phi})|$. Hence there exists an inner function $F_0(z)$ such that $F(z) = F_0(z)F_1(z)$; Beurling [1]. This or the corresponding factorization of the boundary functions $f^+(\phi) = f_0^+(\phi)f_1^+(\phi)$ will be called a factorization of $F(z)$ or $f^+(\phi)$ into inner and outer functions. This factorization is unique up to constant factors of absolute value 1.

Note finally that if $F_1(z)$ and $f_1^+(\phi) = F_1(e^{i\phi})$ are outer functions, then so are the reciprocals $1/F_1(z)$ and $1/f_1^+(\phi)$. Also, for a fixed $p > 0$, $F_1(z) \in H^p$ if and only if $|f_1^+(\phi)| = e^{\alpha(\phi)} \in L^p$.

Proof of Theorem 2.1. Suppose first that $M(f)$ is irreducible. Since $f(\phi) \in M(f) = M(f_0)$, it follows from Lemma 2.1 that $f(\phi)$ has a representation (2.4) where $f_1^+(\phi) \in L^{2+}$. Since $|f^+(\phi)| = |f(\phi)| \in L^2$, it follows that $\log|f^+(\phi)| \in L^1$; Szegö [9] (or the F. Riesz [6] result mentioned above). Hence (2.3) holds. This proves the "only if" assertions.

Suppose that $f(\phi) \in L^2$ satisfies (2.3). Then $|f(\phi)|^2 \in L^1$ and (2.3) imply (Szegö [9], p. 235) that $|f(\phi)|^2$ has a factorization

$$|f(\phi)|^2 = f_1^+(\phi)\bar{f}_1^+(\phi) = |f_1^+(\phi)|^2, \tag{2.8}$$

where $f_1^+(\phi) \in L^{2+}$ is the outer function which is the boundary function of $F_1(z) = \exp A(z)$ and $A(z)$ is given by (2.7), $\alpha(\phi) = \log|f(\phi)|$. Since $|f(\phi)| = |f_1^+(\phi)|$, the factorization (2.4) follows. (This corresponds to the factorization given by Beurling in the case $f \in L^{2+}$.)

Thus in order to complete the proof of the "if" assertions, it suffices to show that when $f(\phi)$ has a factorization (2.4), then $M(f) = f_0(\phi)\cdot L^{2+}$. But this follows from $|f_0(\phi)| = 1$ and the Beurling result: $M(f_1^+) = L^{2+}$ for outer functions $f_1^+(\phi) \in L^{2+}$.

The uniqueness of $f_0(\phi)$ is contained in Lemma 2.1. This proves the theorem.

For later purposes, another representation of the factorization (2.4) will be given. By the theorem of Szegö [9] used above, $|f(\phi)|$ has a factorization

$$|f(\phi)| = f_2^+(\phi)\bar{f}_2^+(\phi) = |f_2^+(\phi)|^2, \tag{2.9}$$

where $f_2^+(\phi)$ is an outer function and, in fact, can be chosen so that $f_1^+(\phi) = [f_2^+(\phi)]^2$. One then obtains the factorization (2.4) with $f_0(\phi)$ given by

$$f_0(\phi) = (f(\phi)/|f(\phi)|)(\bar{f}_2^+(\phi)/f_2^+(\phi)). \tag{2.10}$$

3. The subspaces $N(f)$. Let P denote the orthogonal projection of

L^2 onto L^{2+} so that, in terms of (1.1) and (1.3), $Pf(\phi) = f^+(\phi)$. For $f(\phi) \in L^2$, introduce the subspace

$$N(f) = \text{closure of } PM(f) \tag{3.1}$$

of L^{2+}. There will be derived conditions in order that $N(f) = L^{2+}$. A complete characterization of $N(f)$ (for example, as Lemma 2.1 gives for irreducible $M(f)$) will not be obtained.

Before dealing with the question, a motivation for the consideration of $N(f)$ will be given. For a given $f \in L^2$, $T(f)$ is an operator from L^{2+} to L^{2+}. Its domain $\mathfrak{D}(T)$ is the set of all $x^+(\phi) \in L^{2+}$ such that if $y^+(\phi) + y^-(\phi) = f(\phi)x^+(\phi)$, then $y^+(\phi) \in L^{2+}$ and $y^+(\phi) = T(f)x^+(\phi)$. Roughly speaking, $T(f)x^+(\phi) = Pf(\phi)x^+(\phi)$, where however $f(\phi)x^+(\phi) \in L^1$ is not required to be in L^2. Let $T_0 = T_0(f)$ be the restriction of T with domain $\mathfrak{D}(T_0)$ consisting of trigonometric polynomials $x^+(\phi) \in L^{2+}$. Thus $N(f)$ is the closure of the range of $T_0(f)$. Also $T_0{}^*(f) = T(\bar{f})$, i.e., the adjoint of $T_0(f)$ is $T(\bar{f})$. The adjoint $T^*(\bar{f})$ of $T(\bar{f})$ is the closure of $T_0(f)$. Thus

$$T_0(f) \subset T^*(\bar{f}) \subset T(f), \qquad T^{**}(f) = T(f). \tag{3.2}$$

Since the null space $\mathfrak{N}(T^*)$ of the adjoint T^* of a closed operator T is the orthogonal complement of the closure of the range $\mathfrak{R}(T)$ of T, one has

LEMMA 3.1. *Let $f(\phi) \in L^2$. Then $N(f)^\perp = \mathfrak{N}(T(\bar{f}))$, where $N(f)^\perp$ is the orthogonal complement of $N(f)$ in L^{2+}.*

In particular, if $g(\phi) \in L^2$ is real-valued, then $T^*(g)$ is a closed, symmetric operator and

$$g = \bar{g} \Rightarrow T_0(g) \subset T^*(g) \subset T(g) = T^{**}(g) \tag{3.3}$$

cf. [4]. Thus, the deficiency indices of $T^*(g)$ are the dimensions of $\mathfrak{N}(T(g \pm i))$ or, equivalently, the dimensions of the co-spaces of the closure of the ranges of $T_0{}^*(g \pm i)$, i.e., $\dim N(g \mp i)^\perp$. This gives the following

LEMMA 3.2. *Let $g(\phi) \in L^2(\phi)$ be real-valued. Then $T(g)$ is self-adjoint if and only if $N(g \pm i) = L^{2+}$. (In which case, the operators $T(g \pm i)$ are invertible, i.e., have unique, bounded, left and right inverses.)*

Remark. If $h(\phi) \in L^2$, then $g(\phi) \in L^2$ is orthogonal to $h(\phi)e^{in\phi}$ for $n = 0, \pm 1, \cdots$ if and only if $h(\phi)g(\phi) = 0$. Thus, as observed in Sarason [7], the smallest subspace of L^2 which reduces U and contains $h(\phi)$ (i.e., the subspace of L^2 spanned by $h(\phi)e^{in\phi}$ for $n = 0, \pm 1, \cdots$) consists of the set of functions in L^2 which vanish on the ϕ-set where $h(\phi) = 0$. It has been

pointed out to me by Sarason that this fact implies that *if $M(f)$ is not irreducible, then $N(f) = L^{2+}$.* For suppose that $N(f) \neq L^{2+}$, then there exists an $h^+(\phi) \in L^{2+}$ which is orthogonal to $N(f)$ and, hence, to $M(f)$ and $h^+(\phi) \not\equiv 0$. Let $M_0 \neq \{0\}$ be a subspace of $M(f)$ which reduces U, then $h^+(\phi) \in M_0{}^\perp$, the orthogonal complement of M_0 in L^2. Since $h^+(\phi) \neq 0$ almost everywhere by a classical theorem of F. and M. Riesz, the remark at the beginning of this paragraph implies that the subspace $M_0{}^\perp$ which reduces U and contains $h^+(\phi)$ is L^2. Hence $M_0 = \{0\}$. This contradiction proves the italicized statement.

Thus, in dealing with the subspaces $N(f)$, it will be supposed that $M(f)$ is irreducible. It suffices, therefore, to consider only $N(f_0)$, where $f_0 \in L^{2+}$ and $|f_0(\phi)| = 1$. Whenever convenient, the notation f_0, $f_1{}^+$, $f_2{}^+$ of (2.8)-(2.10) will be used.

4. Conditions for $N(f) = L^{2+}$. The results will be stated in this section. Their proofs will be deferred to the following sections.

THEOREM 4.1. *Let $f(\phi) \in L^2$ and $\log|f(\phi)| \in L^1$. Necessary and sufficient conditions for*

$$e_0(\phi) \equiv 1 \in PM(f), \tag{4.1}$$

(hence, sufficient conditions for $N(f) = L^{2+}$) are the following equivalent conditions:

(i) *there exist inner and outer functions $x_0{}^+(\phi), x_1{}^+(\phi) \in L^{2+}$ such that $f_0(\phi)$ is representable in the form*

$$f_0(\phi) = \bar{x}_0{}^+(\phi)\bar{x}_1{}^+(\phi)/x_1{}^+(\phi); \tag{4.2}$$

(ii) *there exists an inner function $x_0{}^+(\phi) \in L^{2+}$ and a function $\theta(\phi) = \arg[f(\phi)x_0{}^+(\phi)]$ which has a conjugate function $\psi(\phi) \in L^1$ satisfying*

$$e^{\psi(\phi)}|f(\phi)| \in L^1. \tag{4.3}$$

The condition (4.1) *is not necessary for $N(f) = L^{2+}$.*

The validity of the parenthetical assertion following (4.1) is clear; cf. Section 5. In fact, much of the following is based on the fact that $N(f) = L^{2+}$ *if and only if $e_0(\phi) \in N(f)$.*

Note that in (ii), it is not stated that $\theta(\phi) \in L^1$. Condition (ii) should be interpreted in the sense that there exists a real-valued $\psi(\phi) \in L^1$ satisfying (4.3) and having a conjugate function $-\theta(\phi) = -\arg[f(\phi)x_0{}^+(\phi)]$.

An immediate consequence of Theorem 4.1 is the following:

COROLLARY 4.1. *Let* $f(\phi) \in L^2$ *and* $\log|f(\phi)| \in L^1$. *A sufficient condition for* $N(f) = N(\bar{f}) = L^{2+}$ *is that there exists a function* $\theta(\phi) = \arg f(\phi)$ *which has a conjugate function* $\psi(\phi) \in L^1$ *satisfying*

$$e^{|\psi(\phi)|}\,|f(\phi)| \in L^1. \tag{4.4}$$

If, in addition, it is assumed that $1/f(\phi)$ is bounded, then, as will be seen in Section 6, the condition on the existence of $\theta(\phi) = \arg f(\phi)$ in this corollary is necessary and sufficient for $e_0(\phi) \in PM(f) \cap PM(\bar{f})$. This gives the following:

COROLLARY 4.2. *Let* $f(\phi)$ *and* $1/f(\phi)$ *be bounded and measurable functions. Then* $e_0(\phi) \in \mathfrak{R}(T(f)) \cap \mathfrak{R}(T(\bar{f}))$ *if and only if there exists a function* $\theta(\phi) = \arg f(\phi)$ *having a conjugate function* $\psi(\phi) \in L^1$ *satisfying*

$$e^{|\psi(\phi)|} \in L^1. \tag{4.5}$$

In this case, there exists a unique (up to a multiplicative constant) outer function $x_1^+(\phi)$ *such that* x_1^+, $1/x_1^+ \in L^{2+}$ *and* $f(\phi)/|f(\phi)| = \bar{x}_1^+(\phi)/x_1^+(\phi)$.

In Corollary 4.2, the condition (4.4) is necessary for $T(f)$ to be invertible. A necessary and sufficient condition is that the matrix which is the formal product $T(x_1^+)T(1/\bar{x}_1^+)$ be bounded; cf. [4], pp. 871-873; for an interpretation of this matrix product, see Widom [11].

Remark 1. It will follow from Theorem 4.2 that if (4.1) holds, then a necsesary condition for $\lambda = 0$ not to be in the point spectrum of $T(f)$ (i.e., a necessary condition for $N(\bar{f}_0) = L^{2+}$) is that $f_0(\phi)$ have a representation as in (i) with $x_0^+(\phi) \equiv 1$. This condition is not sufficient; cf. Remark 3 and its proof. (Note that although $N(f) = N(f_0)$, it is not necessarily true for unbounded f that $N(\bar{f}) = N(\bar{f}_0)$.)

Remark 2. The relation (4.1) is equivalent to $e_0(\phi) \in \mathfrak{R}(T(f_0))$, the range of $T(f_0)$. A modification of the proof of Theorem 4.1 shows that $e_0(\phi) \in \mathfrak{R}(T(f))$ if and only if (ii) holds with (4.3) replaced by

$$e^{\psi(\phi)}/|f(\phi)| \in L^1. \tag{4.6}$$

The relation $e_0(\phi) \in \mathfrak{R}(T(f))$ is equivalent to a representation of $f_0(\phi)$ in the form (4.2) and of $f(\phi)$ in the form

$$f(\phi)x^+(\phi) = \bar{y}^+(\phi). \tag{4.7}$$

where $x^+(\phi) \in L^{2+}$, $y^+(\phi) \in L^{1+}$. If $\lambda = 0$ is not in the point spectrum of $T(f)$, then $x_0^+(\phi)$ can be chosen to be 1 (i.e., $\theta(\phi) = \arg f(\phi)$) in this necessary and sufficient condition.

The sufficient condition (4.3) for $N(f) = L^{2+}$, but not for (4.1), can be lightened somewhat.

LEMMA 4.1. *Let $f(\phi) \in L^2$ and $\log|f(\phi)| \in L^1$. A sufficient condition for $N(f) = L^{2+}$ is that there exist an inner function $x_0^+(\phi)$ and a $\theta(\phi) = \arg[f(\phi)x_0^+(\phi)]$ having a conjugate function $\psi(\phi) \in L^1$ satisfying*

$$(4.8) \qquad e^{\frac{1}{2}\psi(\phi)}|f(\phi)|^{\frac{1}{2}} \in L^1 \text{ and } e^{p\psi(\phi)}|f(\phi)| \in L^1 \text{ for } p < 1,$$

$$(4.9) \qquad \int_0^{2\pi} |1 - e^{i\epsilon\theta(\phi)}|^2 e^{(1-\epsilon)\psi(\phi)} |f(\phi)| \, d\phi \to 0 \text{ as } \epsilon \to +0.$$

A necessary and sufficient condition for $N(f) = L^{2+}$ can be deduced from the following

THEOREM 4.2. *Let $f(\phi) \in L^2$ and $\log|f(\phi)| \in L^1$. Necessary and sufficient for $N(f) \neq L^{2+}$ (i.e., for $\lambda = 0$ to be in the point spectrum of $T(\bar{f})$) are the following two equivalent conditions:*

(i) *there exists a non-constant inner function $x_0^+(\phi)$ and an outer function $x_1^+(\phi) \in L^{2+}$ such that $f_0(\phi)$ is representable in the form*

$$(4.10) \qquad f_0(\phi) = x_0^+(\phi)(1/\bar{x}_1^+(\phi))/(1/x_1^+(\phi));$$

(ii) *there exists a non-constant inner function $x_0^+(\phi)$ and a $\theta(\phi) = \arg[f(\phi)\bar{x}_0^+(\phi)]$ having a conjugate function $\psi(\phi) \in L^1$ satisfying*

$$(4.11) \qquad e^{-\psi(\phi)}/|f(\phi)| \in L^1.$$

If (i) *and/or* (ii) *holds, $1 + \dim N(f_0)^\perp$ is at least as large as the number of linearly independent divisors of $x_0^+(\phi)$.*

An inner function $y_0^+(\phi)$ is called a divisor of $x_0^+(\phi)$ if there exists an inner function $z_0^+(\phi)$ such that $x_0^+(\phi) = y_0^+(\phi)z_0^+(\phi)$; Beurling [1]. For example, if $x_0^+(\phi) = e^{im\phi}y_0^+(\phi)$, where $m > 0$ and $y_0^+(\phi)$ is an inner function (possibly $e_0(\phi) \equiv 1$), then $1, e^{i\phi}, \cdots, e^{im\phi}$ are divisors of $x_0^+(\phi)$ and $\dim N(f_0)^\perp \geqq m$.

Note that (4.10) is equivalent to a representation for $\bar{f}_0(\phi)$,

$$(4.12) \qquad \bar{f}_0(\phi) = \bar{x}_0^+(\phi)\bar{x}_1^+(\phi)/x_1^+(\phi),$$

analogous to that of (4.2) for $f_0(\phi)$.

In order to state the next theorem, introduce the following notation: For an inner function $x_0^+(\phi)$, let $M_p(x_0^+)$ be the closure in L^p, $1 \leqq p \leqq \infty$, of the linear manifold spanned by $x_0^+(\phi)e^{in\phi}$, $n = 0, 1, \cdots$ and let $M_p(x_0^+)^\perp$ be the orthogonal complement of $M_p(x_0^+)$ in L^{q+}, $1/p + 1/q = 1$.

COROLLARY 4.3. *Let the condition* (i) *of Theorem* 4.2 *hold and, in addition, let* $1/x_1^+(\phi) \in L^2$ (*or equivalently, let* (ii) *and, in addition,* (4.3) *hold*). *Then* $N(f_0)^\perp = \mathfrak{N}(T(\bar{f}))$ *is given by*

$$\mathfrak{N}(T(\bar{f})) = \{x^+(\phi) \in L^{2+}: x^+(\phi)/x_1^+(\phi) \in M_\infty(x_0^+)^\perp\}; \tag{4.13}$$

that is, $\mathfrak{N}(T(\bar{f})) = L^{2+} \cap \{x_1^+(\phi) \cdot M_\infty(x_0^+)^\perp\}$. (*For example, if* $x_0^+(\phi) = e^{im\phi}$, *then* $\dim \mathfrak{N}(T(\bar{f})) = m$.)

Remark 3. This corollary becomes false if the assumption $1/x_1^+(\phi) \in L^2$ or (4.3) is omitted. The difficulty arises from the fact that in the representation (4.10), the outer function $x_1^+(\phi) \in L^{2+}$ is not unique up to constant factors (even if x_0^+ is fixed). (If any representation (4.10) exists, however, with $x_0^+(\phi)$ an inner function and $x_1^+(\phi) \in L^{2+}$ an outer function such that $1/x_1^+(\phi) \in L^{2+}$, then the functions x_0^+, x_1^+ are unique up to constant factors. It will be verified in Section 8 that a necessary condition for $\infty > \dim N(f_0)^\perp = m > 0$ is that $f_0(\phi)$ have a representation (4.10) (i.e., $\bar{f}_0(\phi)$ have a representation (4.12)) with $x_0^+(\phi) = e^{im\phi}$ and $x_1^+(\phi) \in L^{2+}$ on outer function. But this condition is only sufficient for $\dim N(f_0)^\perp \geqq m$, and $\dim N(f_0)^\perp > m$ can hold.

Theorem 4.2 and its proof imply the following:

COROLLARY 4.4. *Let* $f_1(\phi), f_2(\phi) \in L^2$, $\log|f_1(\phi)| \in L^1$, $f_2(\phi)/f_1(\phi)$ *real-valued, and* $f_2(\phi)/f_1(\phi) \geqq c > 0$. *Then* $\dim \mathfrak{N}(T(f_1)) \leqq \dim \mathfrak{N}(T(f_2))$. *In fact, if* $x^+(\phi) \in \mathfrak{N}(T(f_1))$, *then* $x^+(\phi) f_{12}^+(\phi)/f_{22}^+(\phi) \in \mathfrak{N}(T(f_2))$, *where* f_{12}^+, f_{22}^+ *belong to* f_1, f_2, *respectively, as* f_2^+ *belongs to* f *in* (2.9).

This, in turn, will be seen to have the following consequence:

COROLLARY 4.5. *Let* $f_1(\phi), f_2(\phi) \in L^2$, $\log|f_1(\phi)| \in L^1$, *and* $f_2(\phi)/f_1(\phi)$ *real-valued and* $0 < c \leqq f_2(\phi)/f_1(\phi) \leqq C$. *Then* $T(f_1)$ *is invertible if and only if* $T(f_2)$ *is invertible.*

Theorem 4.1 will be proved in the next section and Corollary 4.2 in Section 6. The proof of Lemma 4.1 will be given in Section 7. The next section, Section 8, contains the proof of Theorem 4.2, Corollary 4.3 and the Remark 3. Finally, Corollary 4.5 is proved in Section 9.

5. Proof of Theorem 4.1. Suppose first that $e_0(\phi) \in PM(f_0)$, i.e., there exists a $z^+(\phi) \in L^{2+}$ such that

$$f_0(\phi) z^+(\phi) = \bar{y}^+(\phi), \qquad y^+(\phi) \sim 1 + y_1 e^{i\phi} + \cdots \in L^{2+}. \tag{5.1}$$

Then $z^+(\phi)$ and $y^+(\phi)$ have factorizations $z^+ = z_0^+ z_1^+$, $y^+ = y_0^+ y_1^+$ into inner

and outer functions in L^{2+}. Since $|z_1^+| = |y_1^+|$, the outer functions z_1^+, y_1^+ differ only by a constant factor of absolute value 1. It can be supposed therefore that $z_1^+ = y_1^+$ (by suitably modifying z_0^+ if necessary). Thus (4.2) holds with $x_0^+(\phi) = y_0^+(\phi)/\bar{z}_0(\phi) = y_0^+(\phi)z_0^+(\phi)$ and $x_1^+(\phi) = y_1^+(\phi)$. Hence, $(4.1) \Rightarrow$ (i).

Let (4.2) hold and let

$$x_1^+(\phi) \sim d_0 + d_1 e^{i\phi} + \cdots, \text{ where } d_0 \neq 0. \tag{5.2}$$

Then (5.1) holds if

$$z^+(\phi) = x_0(\phi)x_1^+(\phi)/\bar{d}_0. \tag{5.3}$$

i.e., (i) $\Rightarrow$ (4.1). Thus (4.1) and (i) are equivalent.

From (4.2), it follows that $f_0(\phi)x_1^+(\phi) = \bar{x}_0^+(\phi)\bar{x}_1^+(\phi)$. Hence it is clear that $M(f_0) = M(\bar{x}_0^+\bar{x}_1^+)$ and that $N(f_0) = L^{2+}$. This gives the parenthetical part of the theorem to the effect that (4.1) is sufficient for $N(f) = L^{2+}$.

It will now be shown that (i) $\Rightarrow$ (ii). From (4.2) and (2.10),

$$x_0^+(\phi)f(\phi)/|f(\phi)| = [\bar{x}_1^+(\phi)/\bar{f}_2^+(\phi)]/[x_1^+(\phi)/f_2^+(\phi)]. \tag{5.4}$$

Since $x_1^+(\phi)/f_2^+(\phi)$ is an outer function, there is a real-valued function $\psi(\phi) \in L^1$ having a conjugate function $-\theta(\phi)$ such that

$$x_1^+(\phi)/f_2^+(\phi) = e^{\frac{1}{2}[\psi - i\theta]}, \tag{5.5}$$

so that $\theta(\phi) = \arg[f(\phi)x_0^+(\phi)]$. Also $|x_1^+(\phi)|^2 = e^{\psi}|f_2^+(\phi)|^2 = e^{\psi}|f(\phi)|$ is of class L^1. Consequently, (i) $\Rightarrow$ (ii).

Conversely, if (ii) is assumed, then

$$x_0^+(\phi)f(\phi)/|f(\phi)| = \bar{y}_1^+(\phi)/y_1^+(\phi), \tag{5.6}$$

where $y_1^+ = e^{\frac{1}{2}(\psi - i\theta)}$ is an outer function such that $|y_1^+|^2|f| \in L^1$. Hence $x_1^+ = y_1^+f_2^+$ is an outer function of class L^{2+}. This gives (5.4) and hence (4.2) by (2.10).

In order to see that (4.1) is not necessary for $N(f_0) = L^{2+}$, let $f_0(\phi)$ be a real-valued function assuming the values ± 1 but is not a constant almost everywhere. Then [4, (iii)] implies that $e_0(\phi) \notin N(f_0)$, i.e., (i), (ii) do not hold. But $N(f_0) = L^{2+}$ for otherwise, $\lambda = 0$ is in the point spectrum of $T(\bar{f}_0)$; Lemma 3.1. But, since $f_0 = \bar{f}_0$ is bounded, the point spectrum of $T(f_0)$ is empty by [4, (ii)]. This completes the proof of the theorem.

6. Proof of Corollary 4.2. Instead of proving the corollary directly, the more general statement preceding it will first be proved. That is, if

$f(\phi) \in L^2$ and $1/f(\phi)$ is bounded, then $e_0(\phi) \in PM(f) \cap PM(\bar{f})$ if and only if there exists a $\theta(\phi) = \arg f(\phi)$ having a conjugate function $\psi(\phi) \in L^1$ satisfying (4.4). (The uniqueness statement at the end of Corollary 4.2 will follow from the proof of this statement.)

By the proof of Theorem 4.1, $e_0(\phi) \in PM(f)$ if and only if there exist inner and outer functions $x_0^+, x_1^+ \in L^{2+}$ such that (5.4) holds. Similarly, $e_0(\phi) \in PM(\bar{f})$ if and only if there exist inner and outer functions $y_0^+, y_1^+ \in L^{2+}$ such that

$$(6.1) \qquad y_0^+(\phi)\bar{f}(\phi)/|f(\phi)| = [\bar{y}_1^+(\phi)/\bar{f}_2^+(\phi)]/[y_1^+(\phi)/f_2^+(\phi)].$$

When (5.4) and (6.1) hold, a multiplication of these relations gives

$$(6.2) \qquad z_0^+(\phi)z_1^+(\phi) = \bar{z}_1^+(\phi),$$

where

$$(6.3) \qquad z_0^+ = x_0^+y_0^+ \text{ and } z_1^+ = (x_1^+/f_2^+)(y_1^+/f_2^+)$$

are, respectively, an inner function and an outer function of class L^{1+}, since $1/f_2^+$ is bounded. It follows that both sides of (6.2) are constants and, since the only inner function which "divides" a constant is a constant, x_0^+ and y_0^+ are constants. Thus in (5.4) and (6.1), it can be supposed that $x_0^+ \equiv 1$, $y_0^+ \equiv 1$ and $z_1^+ \equiv 1$. Thus if (5.5) holds, then

$$(6.4) \qquad y_1^+(\phi)/f_2^+(\phi) = e^{-\frac{1}{2}[\psi - i\theta]}.$$

Consequently, $e^{\pm\psi}|f| \in L^1$, i.e., $e^{|\psi|}|f| \in L^1$.

Corollary 4.2 follows from the fact that $PM(f) = \mathcal{R}(T(f_0))$ and that, when f and $1/f$ are bounded, $\mathcal{R}(T(f_0)) = \mathcal{R}(T(f))$. For example, $T(f_0)x^+ = y^+$ if and only if $T(f)[x^+/f_1^+] = y^+$. This completes the proof.

7. Proof of Lemma 4.1. The assumptions of the lemma show that

$$(7.1) \qquad e^{i\theta(\phi)} = (f(\phi)/|f(\phi)|)x_0^+(\phi) = \bar{x}^+(\phi)/x^+(\phi),$$

where $x^+(\phi) = e^{\frac{1}{2}(\psi - i\theta)}$ is an outer function. Hence, by (2.10),

$$(7.2) \qquad f_0(\phi) = \bar{x}_0^+(\phi)\bar{x}^+(\phi)\bar{f}_2^+(\phi)/x^+(\phi)f_2^+(\phi).$$

Since $|f_2^+(\phi)| = |f(\phi)|^{\frac{1}{2}}$, the first assumption in (4.8) implies that $x^+(\phi)f_2^+(\phi)$ is an outer function of class L^{1+}. In particular,

$$(7.3) \qquad (2\pi)^{-1}\int_0^{2\pi} x^+(\phi)f_2^+(\phi)\,d\phi = c_0 \text{ exists and } \neq 0.$$

For $0 < \epsilon < 1$, put

$$(7.4) \qquad c_\epsilon = (2\pi)^{-1} \int_0^{2\pi} (x^+(\phi))^{1-\epsilon} f_2^+(\phi)\, d\phi.$$

Note that the outer function $x^+(\phi)$ is the boundary function of an outer function $F(z)$ which is regular and does not vanish for $|z| < 1$. Hence $(x^+(\phi))^\epsilon$ can be taken as the boundary function of $(F(z))^\epsilon$ which is uniquely detremined up to a factor $e^{2\pi i n/\epsilon}$ independent of z. It will be supposed that this determination is chosen so that

$$(7.5) \qquad e^{i\epsilon\theta(\phi)} = (\bar{x}^+(\phi))^\epsilon / (x^+(\phi))^\epsilon$$

where $(\bar{x}^+(\phi))^\epsilon$ is the complex-conjugate of $(x^+(\phi))^\epsilon$. Also, $(x^+(\phi))^{1-\epsilon}$ in (7.4) and below is defined as $x^+(\phi)/(x^+(\phi))^\epsilon$.

The integral in (7.4) exists. In fact, the absolute value of the integrand is $e^{\frac{1}{2}(1-\epsilon)\psi} | f_2^+ |$ which is majorized by $(e^{\frac{1}{2}\psi} + 1) | f |^{\frac{1}{2}} \in L^1$. By Lebesgue's theorem on dominated convergence,

$$(7.6) \qquad c_\epsilon \to c_0 (\neq 0) \text{ as } \epsilon \to +0.$$

Define $f_\epsilon(\phi)$ by

$$(7.7) \qquad f_\epsilon(\phi) = f_0(\phi) e^{-i\epsilon\theta(\phi)}.$$

Hence (7.2) and (7.4) show that

$$(7.8) \qquad f_\epsilon(\phi) y_\epsilon(\phi) = (\bar{x}^+(\phi))^{1-\epsilon} \bar{f}_2^+(\phi) / \bar{c}_\epsilon \sim 1 + \cdots,$$

where

$$(7.9) \qquad y_\epsilon(\phi) = x_0^+(\phi) (x^+(\phi))^{1-\epsilon} f_2^+(\phi) / \bar{c}_\epsilon$$

is of class L^{2+} by (4.8). The relation (7.8) gives $T(f_\epsilon) y_\epsilon = e_0(\phi)$. Hence

$$\| T(f_0) y_\epsilon - e_0(\phi) \| = \| (T(f_0) - T(f_\epsilon)) y_\epsilon \| \leqq \| (f_0 - f_\epsilon) y_\epsilon \|.$$

By (7.6), (7.7) and (7.9), the condition (4.9) means that

$$\| (f_0 - f_\epsilon) y_\epsilon \|^2 \to 0 \text{ as } \epsilon \to 0.$$

Hence $e_0(\phi) \in N(f)$, the closure of the range of $T(f_0)$. Consequently, $N(f) = L^{2+}$.

8. Proof of Theorem 4.2. Since the proof of the equivalence of (i) and (ii) in Theorem 4.2 is the same as in Theorem 4.1, only (i) will be considered.

It is easy to see that $0 \neq v^+(\phi) \in L^{2+}$ is in $N(f_0)\perp$ if and only if $\bar{f}_0 v^+$ is orthogonal to L^{2+}, i.e.,

$$\bar{f}_0(\phi)v^+(\phi) = e^{-i\phi}\bar{w}(\phi), \text{ where } w^+(\phi) \in L^{2+}. \tag{8.1}$$

Arguing as in Section 5, one obtains from (8.1) a representation of the form (4.10) in which $x_0^+(\phi)$ has the divisor $e^{i\phi}$.

Conversely, let $\bar{f}_0(\phi)$ have a representation (4.12) with $x_0(\phi) \not\equiv \text{const.}$, let (5.2) hold, and $x_0^+(\phi) \sim c_m e^{im\phi} + \cdots$ with $c_m \neq 0$, $m \geqq 0$. If $m > 0$, then $\lambda = 0$ is in the point spectrum of $T(\bar{f}_0)$; in fact $T(\bar{f}_0)x_1^+(\phi) = 0$. If $m = 0$, then

$$\bar{f}_0(\phi)z^+(\phi) = \bar{y}^+(\phi), y^+(\phi) \sim 1 + y_1 e^{i\theta} + \cdots \in L^{2+}$$

holds with $z^+(\phi) = x_1^+(\phi)/\bar{c}_0\bar{d}_0$ and also with (5.3). Thus $T(\bar{f}_0)z^+ = 1$ has two linearly independent solutions, so that $\lambda = 0$ is in the point spectrum of $T(\bar{f}_0)$.

If y_0^+, z_0^+ are divisors of x_0^+ and $x_0^+ = y_0^+ z_0^+$, then (4.12) implies $\bar{f}_0 y_0^+ x_1^+ = \bar{z}_0^+ \bar{x}_1^+$. This makes it clear that if x_0^+ has m linearly independent divisors, $1 < m \leqq \infty$, then (8.1) has at least $m-1$ linearly independent solutions $v^+(\phi)$.

Proof of Corollary 4.3. If $x^+(\phi) \in L^{2+}$ is in $N(f_0)\perp$, then the scalar products $(f_0 p^+, x^+)$ vanish for all trigonometric polynomials $p^+(\phi) \in L^{2+}$. Since $1/x_1^+(\phi) \in L^{2+}$ is an outer function, $M(f_0) = M(f_0/x_1^+)$ and $(f_0 p^+, x^+) = 0$ for all p^+ if and only if $((f_0/x_1^+)p^+, x^+) = 0$ for all p^+. In view of (4.10), this is equivalent to $(x_0^+ p^+, x^+/x_1^+) = 0$ for all p^+. Hence $x^+ \in L^2$ is in $N(f_0)\perp$ if and only if $x^+/x_1^+ \in M_\infty(x_0^+)\perp$.

Proof of Remark 3. If $0 \neq v^+(\phi) \in N(f_0)\perp$, i.e., if (8.1) holds, let

$$e^{i\phi}w^+(\phi) \sim w_k e^{ik\phi} + w_{k+1}e^{i(k+1)\phi} + \cdots, \text{ where } w_k \neq 0$$

and $k > 0$ is an integer. The integer k will be said to belong to the element $v^+(\phi) \in N(f_0)\perp$. If $\dim N(f_0)\perp \geqq m > 0$, then there is an element $v^+(\phi) \in N(f_0)\perp$ such that m belongs to $v^+(\phi)$. In this case, (8.1) leads to a representation for $f_0(\phi)$ of the form

$$\bar{f}_0(\phi) = e^{-im\phi}\bar{x}_0^+(\phi)\bar{x}_1^+(\phi)/x_1^+(\phi), \tag{8.2}$$

where x_0^+, x_1^+ are inner, outer functions in L^{2+}. Arguing as in the proof of Theorem 4.2, $x_0^+(\phi) \not\equiv \text{const.}$ implies that $\dim N(f_0)\perp > m$; i.e., if $\dim N(f_0)\perp = m > 0$, then $x_0^+(\phi) \equiv \text{const.}$ This proves the first part of Remark 3.

In order to prove the last part note that if

$$(8.3)\qquad x_1^+(\phi) = i + e^{i\theta}, \qquad y_1^+ = 1 + ie^{i\theta},$$

then x_1^+, y_1^+ are linearly independent, bounded, outer functions and, as is easily verified, $\bar{x}_1^+/x_1 = \bar{y}_1^+/y_1^+$. For a given $m \geqq 0$, let

$$\bar{f}_0(\phi) = e^{-im\phi}\bar{x}_1^+/x_1^+ = e^{-im\phi}\bar{y}_1^+/y_1^+.$$

Thus $\bar{f}_0(\phi)$ has representations of the form (4.12) with $x_0^+(\phi) = e^{im\phi}$. But if $m > 0$, then $\dim N(f_0)^\perp > m$ for it contains the $2m$ elements $v^+ = e^{ik\phi}x_1^+$ and $v^+ = e^{ik\phi}y_1^+$ for $k = 0, 1, \cdots, m-1$. Also if $m = 0$, $\dim N(f_0)^\perp > 0$ for it contains $v^+ = y_1^+ - ix_1^+$. Actually, $\dim N(f_0)^\perp = \infty$ as can be seen by using the formula

$$\bar{f}_0(\phi) = e^{-im\phi}\bar{z}_1^+(\phi)/z_1^+(\phi),$$

where $z_1^+(\phi) = (x_1^+(\phi))^\alpha(y_1^+(\phi))^{1-\alpha}$, $0 \leqq \alpha \leqq 1$, and $(x_1^+(\phi))^\alpha$, $(y_1^+(\phi))^{1-\alpha}$ are the boundary functions of $(i+z)^\alpha$, $(1+iz)^{1-\alpha}$ which reduce to $e^{\frac{1}{2}i\pi\alpha}$, 1, respectively, at $z = 0$.

9. Proof of Corollary 4.5. For any $f \in L^2$, $T(f)$ is a closed operator since it is the adjoint of $T_0(\bar{f})$. Thus, $T(f)$ is invertible if and only if it is one-to-one (i.e., $\dim \mathfrak{N}(T(f)) = 0$)) and onto (i.e., $\mathfrak{R}(T(f)) = L^{2+}$).

Corollary 4.4 shows that $T(f_1)$ is one-to-one if and only if $T(f_2)$ is one-to-one. Thus, the corollary will be proved if it is shown that if $T(f_1)$, $T(f_2)$ are one-to-one and $T(f_1)$ is invertible, then $T(f_2)$ is invertible. This, to be proved under the assumption that there is a function $h(\phi)$ such that

$$(9.1)\qquad f_2(\phi) = h(\alpha)f_1(\phi),$$

$0 < c \leqq h(\phi) \leqq C$.

This proof will depend on the following simple lemma.

LEMMA 9.1. *Let $\sigma^+(\phi) \in L^{p+}$ and $f(\phi) \in L^q$, where*

$$(9.2)\qquad p^{-1} + q^{-1} + 2^{-1} \leqq 1.$$

Then

$$(9.3)\qquad T(f\sigma^+) \supset T(f)T(\sigma^+) \text{ and } T(\bar{\sigma}^+f) \supset T(\bar{\sigma}^+)T(f).$$

Cf. Widom, [11], p. 94, where $p = q = \infty$. The first of the relations is clear. The second follows from the fact that if $\tau^+(\phi) \in L^{r+}$, where $p^{-1} + r^{-1} \leqq 1$, then $\sigma^+(\phi)\tau^+(\phi) \in L^{1+}$.

There exists an outer function $h_2^+(\phi)$ such that h_2^+, $1/h_2^+ \in L^{\infty+}$ and

$$h(\phi) = \bar{h}_2^+(\phi) h_2^+(\phi); \tag{9.4}$$

cf. (2.9). Hence

$$f_2(\phi) = \bar{h}_2^+(\phi) f_1(\phi) h_2^+(\phi) \tag{9.5}$$

and, by Lemma 9.1,

$$T(f_2) \supset T(\bar{h}_2^+) T(f_1) T(h_2^+). \tag{9.6}$$

The product on the right of (9.6) has the bounded inverse $T(1/h_2^+)[T(f_1)]^{-1}$ $T(1/\bar{h}_2^+)$. In particular, its range is L^{2+}. Since $T(f_2)$ is one-to-one, it follows that "$\supset$" can be replaced by "$=$" in (9.6). Hence $T(f_2)$ is invertible.

10. Self-adjoint $T(g)$. In view of Lemma 3.2, the results of Section 4 applied to $f = g \pm i$ give criteria for the Toeplitz matrix $T(g)$ of a real-valued $g(\phi) \in L^2$ to be or not to be self-adjoint. These results can be summarized as follows:

Let $g(\phi) \in L^2$ be real-valued. Then

(*) $T(g)$ is self-adjoint if there exists a $\theta(\phi) = \arctan 1/g(\phi)$ having a conjugate function $\psi(\phi) \in L^1$ satisfying

$$e^{|\psi(\phi)|}(1 + |g(\phi)|) \in L^1; \tag{10.1}$$

(**) the deficiency index $\dim \mathfrak{N}(T(g - i)) = 0$ if there exists an inner function $x_0^+(\phi)$ and a $\theta(\phi) = \arctan 1/g(\phi) + \arg x_0^+(\phi)$ having a conjugate function $\psi(\phi) \in L^1$ satisfying

$$e^{\psi(\phi)}(1 + |g(\phi)|) \in L^1 \tag{10.2}$$

or more generally,

$$e^{(1-\epsilon)\psi(\phi)}(1 + |g(\phi)|) \in L^1 \text{ for } 0 < \epsilon < 1, \tag{10.3}$$

$$\int_0^{2\pi} |1 - e^{i\epsilon\theta}|^2 e^{(1-\epsilon)\psi}(1 + |g|)\, d\phi \to 0 \text{ as } \epsilon \to 0. \tag{10.4}$$

If (10.2) holds and, in addition, $x_0^+(\phi) \neq \text{const.}$, then $T(g)$ is maximal symmetric with $\dim \mathfrak{N}(T(g + i)) > 0$;

(***) the deficiency index $\dim \mathfrak{N}(T(g - i)) > 0$ if and only if there exists a non-constant inner function $x_0^+(\phi)$ and a $\theta(\phi) = \arctan 1/g(\phi) - \arg x_0^+(\phi)$ having a conjugate function $\psi(\phi) \in L^1$ satisfying

$$e^{-\psi(\phi)}/(1 + |g(\phi)|) \in L^1; \tag{10.5}$$

in which case, $\dim \mathfrak{N}(T(g-i)) \geqq m-1$ if there exists at least m linearly independent inner functions which divide $x_0^+(\phi)$. (If, in addition, (10.2) holds, then Corollary 4.3 is applicable to $f=g(\phi)+i$.)

(****) $T(g)$ is self-adjoint if and only if $T(gh)$ is self-adjoint, where $h(\phi)$ is a real-valued measurable function such that $0<c\leqq h(\phi)\leqq C$.

In order to verify (****), note that the dimensions m, n of the null spaces of $T(gh\pm i)$ are the same as those of $T(g\pm i/h)=T(g)\pm iT(1/h)$; Corollary 4.4. Since $T(1/h)=[T(h_2^+)T^*(h_2^+)]^{-1}$ in the notation of (9.4), m and n are the dimensions of the null spaces of $S^*=T^*(h_2^+)T(g)T(h_2^+)\pm i$. Thus m, n are the deficiency indices of the closed, symmetric operator $S=T^*(h_2^+)T^*(g)T(h_2^+)$. Since these deficiency indices are the dimensions of the largest subspaces modulo $\mathfrak{D}(S)$ on which $\operatorname{Im}(S^*x^+,x^+)\geqq 0$ or $\leqq 0$ (cf. [5], p. 149) and $(S^*x^+,x^+)=(T(g)T(h_2^+)x^+,T(h_2^+)x^+)$, it is seen that these deficiency indices are the same as those of $T^*(g)$.

In the remainder of this section, illustrations of the use of (*)-(***) will be given.

(a) It was pointed out in [4] that if $g(\phi)\in L^2$ is half-bounded, then $T(g)$ is self-adjoint. It will be indicated how this also follows from (*). Without loss of generality, it can be supposed that $g(\phi)\geqq c>0$, where c can be chosen so that $\theta(\phi)=\arctan 1/g(\phi)$ satisfies $0<\theta(\phi)\leqq\frac{1}{2}$. Then the conjugate function $\psi(\phi)$ is such that $e^{|\psi|}\in L^2$; [12], p. 164, Exercise 3. Since $g\in L^2$, (10.1) holds and hence, $T(g)$ is self-adjoint.

(b) In order to obtain examples of $g(\phi)\in L^2$ which are not half-bounded but for which $T(g)$ is self-adjoint, it suffices to choose $g(\phi)\in L^2$, so that $\theta(\phi)=\arctan 1/g(\phi)$ has a continuous determination. Then $\theta(\phi)$ has a conjugate function $\psi(\phi)$ such that $e^{|\psi|}\in L^p$ for all p; [10], p. 358. Thus (10.1) holds and $T(g)$ is self-adjoint. For example, let $g(\phi)$ be periodic of period 2π and continuous except at a finite number of points $\phi=\phi_1\cdots,\phi_k$ mod 2π and $g(\phi_j+0)=g(\phi_j-0)$ is $\pm\infty$ for $j=1,\cdots,k$.

(c) Let $g(\phi)\in L^2$ be real-valued and extended as a periodic function of period 2π. Let there exist a finite number of points $\phi_1,\cdots,\phi_k$, $\phi_1^+,\cdots,\phi_m^+$, $\phi_1^-,\cdots,\phi_n^-$ mod 2π such that $g(\phi)$ is bounded outside of neighborhoods of these points; $g(\phi)$ is continuous on punctured neighborhoods of these points; both limits $g(\phi\pm 0)$ exist, possibly $\pm\infty$, at these points; one of the limits $g(\phi_j\pm 0)$ is finite for $j=1,\cdots,k$; $g(\phi)$ jumps from $-\infty$ to $+\infty$ at $\phi=\phi_j^+$ for $j=1,\cdots,m$ and jumps from ∞ to $-\infty$ at $\phi=\phi_j^-$ for $j=1,\cdots,n$. It will be shown that (i) *if* $m=0$, *then* $\dim\mathfrak{N}(T(g-i))=0$; *hence* (ii)

if $m = n = 0$, *then* $T(g)$ *is self-adjoint*; (iii) *if* $m \geqq r > 0$ *and if for some* $p > 1$, *the function* $1/|g(\phi)(\phi - \zeta)|^p$ *is integrable for* ϕ *near* $\zeta = \phi_j^-$, $j = 1, \cdots, n$, *and* $\zeta = \phi_j^+$ *for* r *choices of* j, $1 \leqq j \leqq m$, *then* $T(g)$ *is not self-adjoint and* $\dim \mathfrak{N}(T(g - i)) \geqq r > 0$; (iv) *if* $1/g(\phi)$ *is uniformly Dini continuous on every small interval* $[\xi - \epsilon, \xi)$, $(\eta, \eta + \epsilon]$ *where* $|g(\xi - 0)| = \infty$ *and* $|g(\eta + 0)| = \infty$, *then* $\dim \mathfrak{N}(T(g - i)) = m$ *and* $\dim \mathfrak{N}(T(g + i)) = n$.

A function is called uniformly Dini continuous an an interval if it has a degree of continuity $\omega(\delta)$ such that $\int_{+0} \delta^{-1}\omega(\delta)\, d\delta < \infty$.

The proofs of (i)-(iv) depend on a device of Widom [11]. It will remain undecided whether or not $T(g)$ can be self-adjoint if either $m > 0$ or $n > 0$.

In order to prove (i) and (iv), let $\theta(\phi) = \arctan 1/g(\phi)$, $0 \leqq \theta \leqq \pi$. Note that if $g_0(\phi)$ is a bounded, measurable, real-valued function, then $T(g)$ and $T(g + g_0) = T(g) + T(g_0)$ have the same deficiency indices since $T(g_0)$ is bounded and self-adjoint; cf. [5], p. 150. Thus, it can be supposed that $g(\phi)$ is continuous except at $\phi = \phi_j, \phi_j^{\pm}$ and that if $2\pi\alpha_j$ is the jump of $\theta(\phi)$ at ϕ_j, then $4|\alpha_j| < 1$. For otherwise, g is replaced by $g + g_0$ for a suitable choice of g_0.

Let $\gamma(\phi) = \phi - 2\pi[\phi/2\pi]$, where $[x]$ is the integral part of x, denote the continuous periodic function which is ϕ on $0 < \phi < 2\pi$ and has a jump of -2π at $\phi = 0 \bmod 2\pi$. Then

$$\gamma(\phi) \sim \pi - i \sum_{n \neq 0} e^{in\phi}/n,$$

so that its conjugate function is $-2 \log |1 - e^{i\phi}|$ and

$$u(\phi) = \theta(\phi) + \sum_{j=1}^{k} \alpha_j \gamma(\phi - \phi_j) - \tfrac{1}{2} \sum_{j=1}^{m} \gamma(\phi - \phi_j^+) + \tfrac{1}{2} \sum_{j=1}^{n} \gamma(\phi - \phi_j^-) \tag{10.6}$$

is continuous. Consequently, if $v(\phi)$ is the conjugate function of $u(\phi)$, then $e^{|v(\phi)|} \in L^p$ for all p. (Also, if $u(\phi)$ is, for example, uniformly Dini continuous, then $v(\phi)$ is bounded.)

On (i). Suppose that $m = 0$. Then the conjugate function of $\theta(\phi) = \arctan 1/g(\phi)$ is

$$\psi(\phi) = v(\phi) + 2 \sum_{j=1}^{k} \alpha_j \log |1 - e^{i(\phi - \phi_j)}| + \sum_{j=1}^{n} \log |1 - e^{i(\phi - \phi_j^-)}| + C$$

for some constant C. Hence

$$e^{\psi(\phi)} = e^C e^{v(\phi)} \prod_{j=1}^{k} |1 - e^{i(\phi - \phi_j)}|^{2\alpha_j} \prod_{j=1}^{n} |1 - e^{i(\phi - \phi_j^-)}|.$$

The last product is bounded. Since $2\,|\,\alpha_j\,| < \frac{1}{2}$ and $e^{|v|} \in L^p$ for all p, it follows that $e^{\psi} \in L^2$. Thus (10.2) holds when $x_0^+(\phi) \equiv 1$ and so, (i) follows from (**).

On (iii). Assume the conditions of (iii) and suppose that the enumeration of the ϕ_j^+ is chosen so that the first r points $\phi_1^+, \cdots, \phi_r^+$ have the specified property. Write (10.6) as

$$
\begin{aligned}
u(\phi) = \theta_0(\phi) - r\phi + \sum_{j=1}^{k} \alpha_j \gamma(\phi - \phi_j) - \tfrac{1}{2} \sum_{j=r+1}^{m} \gamma(\phi - \phi_j^+) \\
+ \tfrac{1}{2} \sum_{j=1}^{r+m} \gamma(\phi - \zeta_j) + \sum_{j=1}^{r} \phi_j^+,
\end{aligned} \tag{10.7}
$$

where $\zeta_1, \cdots, \zeta_{r+m}$ is the set of points $\phi_1^+, \cdots, \phi_r^+$, $\phi_1^-, \cdots, \phi_m^-$ and $\theta_0(\phi) - \theta(\phi) = 0 \bmod 2\pi$ is the sum of $2\pi[(\phi - \phi_j^+)/2\pi]$ for $j = 1, \cdots, r$. The relation (10.7) is obtained by writing the factor $-\frac{1}{2}$ of $\gamma(\phi - \phi_j^+)$ in (10.6) as $-1 + \frac{1}{2}$ for $j = 1, \cdots, r$; cf. [11], p. 95. Then the conjugate function of $\theta_0(\phi) - r\phi = \arctan 1/g(\phi) - \arg e^{ir\phi}$ is

$$
\begin{aligned}
\psi(\phi) = v(\phi) + 2 \sum_{j=1}^{k} \alpha_j \log |\, 1 - e^{i(\phi - \phi_j)} \,| - \sum_{j=r+1}^{m} \log |\, 1 - e^{i(\phi - \phi_j^+)} \,| \\
+ \sum_{j=1}^{r+m} \log |\, 1 - e^{i(\phi - \zeta_j)} \,| + C.
\end{aligned}
$$

Hence

$$
e^{-\psi(\phi)} = e^{-C} e^{-v(\phi)} \prod_{j=1}^{k} |\, 1 - e^{i(\phi - \phi_j)} \,|^{-2\alpha_j} \prod_{j=r+1}^{m} |\, 1 - e^{i(\phi - \phi_j^+)} \,| / \prod_{j=1}^{r+m} |\, 1 - e^{i(\phi - \zeta_j)} \,|.
$$

Since $e^{-v(\phi)} \in L^p$ for all p and $2\,|\,\alpha_j\,| < \frac{1}{2} < 1$, the assumption on $g(\phi)$ in the neighborhood of $\zeta_j = \phi_1^+, \cdots, \phi_r^+$, $\phi_1^-, \cdots, \phi_m^-$ implies that (10.5) holds. Thus (iii) is a consequence of (***) where $x_0^+(\alpha) = e^{ir\phi}$.

On (iv). It is clear from the proof that, in (iii), "some $p > 1$" can be relaxed to "$p = 1$" if $v(\phi)$ is bounded. This fact will be used in the proof of (iv).

The function $\theta(\phi) = \arctan 1/g(\phi)$, $0 \leqq \theta \leqq \pi$, is uniformly Dini continuous on the intervals occurring in the statement of (iv). Thus, after the addition of a suitable bounded function to $g(\phi)$, it can be supposed that $u(\phi)$ in (10.6) uniformly Dini continuous. Hence its conjugate $v(\phi)$ is continuous and, a fortiori, bounded.

If $\omega(\delta)$ is a degree of continuity for $1/g(\phi)$ on one of the specified intervals, say $[\xi - \epsilon, \xi)$, then $|\, 1/g(\phi) \,| \leqq \omega(\xi - \phi)$ for $\xi - \epsilon \leqq \phi < \xi$. Hence

$$
\int_{\xi-\epsilon}^{\xi} d\phi / |\, g(\phi)(\phi - \xi) \,| \leqq \int_0^{\epsilon} \delta^{-1} \omega(\delta)\, d\delta < \infty.
$$

Similarly, $1/|\, g(\phi)(\phi - \eta) \,|$ is integrable over the interval $(\eta, \eta + \epsilon]$.

Thus $\dim \mathfrak{N}(T(g-i)) \geqq m$ follows from the proof of (iii) and the remark at the beginning of this proof. In fact, $\dim \mathfrak{N}(T(g-i)) = m$ is a consequence of Corollary 4.3 where the analogue of (4.12) holds for $f = g(\phi) + i$ with $x_0^+(\phi) = e^{im\phi}$ and, up to a constant factor,

$$x_1^+(\phi) = e^{-\frac{1}{2}(v-iu)} \prod_{j=1}^{K} (1 - e^{i(\phi-\phi_j)})^{-\alpha_j} \prod_{j=1}^{m+n} (1 - e^{i(\phi-\zeta_j)})^{-\frac{1}{2}} / f_2^+,$$

where ζ_j runs through $\phi_j^\pm$; so that $x_1^+, 1/x_1^+ \in L^{2+}$.

Addendum. C. R. Putnam [Pacific Journal of Mathematics, vol. 9 (1959), p. 838] proved the following: (I) *Let $g(\phi) \in L^2$ be real-valued and even, so that $g(\phi) \sim \sum g_n e^{in\phi} = g_0 + 2 \sum g_n \cos n\phi$, where $g_{-n} = g_n$ is real. If the Hankel matrix $H(g) = (g_{n+m+1})$, where $n, m = 0, 1, 2, \cdots$, is bounded, then $T(g)$ is self-adjoint.* Putnam has pointed out to me that, in view of Nehari's necessary and sufficient condition for an infinite Hankel matrix to be bounded [Annals of Mathematics, vol. 65 (1957), pp. 153-162], (I) takes the following form: (II) *Let $g(\phi)$ be as in (I). If there exists a function $h^+(\phi) \in L^{2+}$ such that $g^+(\phi) + \bar{h}^+(\phi)$ is bounded, then $T(g)$ is self-adjoint.*

It can be mentioned that Putnam's observation (I) has a very simple proof: If $g(\phi) \in L^2$ is real-valued and $L(g)$ is the corresponding Laurent matrix, then, schematically,

$$L(g) = \begin{pmatrix} T(g) & H(g) \\ \bar{H}(g) & T(g) \end{pmatrix},$$

where $\bar{H}(g)$ is the complex conjugate of $H(g)$. If $g(\phi)$ is even, then $H(g)$ is a real matrix and $H(g) = \bar{H}(g)$. If, in addition, $H(g)$ is bounded, then

$$L(g) - \begin{pmatrix} 0 & H(g) \\ H(g) & 0 \end{pmatrix} = \begin{pmatrix} T(g) & 0 \\ 0 & T(g) \end{pmatrix}$$

is self-adjoint since $L(g)$ is. Hence $T(g)$ is self-adjoint.

Added 2/8/63. Professor Devinatz has sent me a copy of his paper "Toeplitz operators on H^2 spaces," to appear, dealing with the invertibility of bounded Toeplitz operators. This paper also employs Beurling factorizations for functions of class H^p. In addition to other results, Devinatz generalizes and proves a theorem announced by H. Widom, Amer. Math. Soc. Notices, vol. 7 (1960), p. 63, which I have overlooked and which gives necessary and sufficient conditions for the invertibility of a bounded $T(f)$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. Beurling, "On two problems concerning linear transformations in Hilbert space," *Acta Mathematica*, vol. 81 (1949), pp. 239-255.

[2] P. R. Halmos, "Shifts on Hilbert spaces," *Journal für die reine und angewandte Mathematik*, vol. 208 (1961), pp. 102-112.

[3] P. Hartman, "Multiplicative sequences and Töplerian (L^2)-bases," *Duke Mathematical Journal*, vol. 14 (1947), pp. 755-767.

[4] ——— and A. Wintner, "The spectra of Toeplitz's matrices," *American Journal of Mathematics*, vol. 76 (1954), pp. 867-882.

[5] M. A. Neumark, *Lineare Differentialoperatoren*, Berlin, (1960).

[6] F. Riesz, "Ueber die Randwerte einer analytischen Funktion," *Mathematische Zeitschrift*, vol. 18 (1923), pp. 87-95.

[7] D. E. Sarason, *The theorem of F. and M. Riesz on the absolute continuity of analytic measures*, Thesis, University of Michigan (1963).

[8] G. Szegő, "Beiträge zur Theorie der Toeplitzschen Formen. I," *Mathematische Zeitschrift*, vol. 6 (1920), pp. 167-202.

[9] ———, "Ueber die Randwerte einer analytischen Funktion," *Mathematische Annalen*, vol. 84 (1931), pp. 232-244.

[10] S. Warschawski, "Ueber einige Konvergenzsätze aus der Theorie der Konformen Abbildung," *Göttinger Nachrichten* (1930), pp. 344-369.

[11] H. Widom, "Inversion of Toeplitz matrices II," *Illinois Journal of Mathematics*, vol. 4 (1960), pp. 88-99.

[12] A. Zygmund, *Trigonometrical Series*, Warsaw-Lwow (1935).

ON STABILITY OF COMPACT SUBMANIFOLDS OF COMPLEX MANIFOLDS.* [1]

By K. KODAIRA.

There are two aspects of stability of compact submanifolds of complex manifolds. Let V be a compact complex analytic submanifold of a complex manifold W. We shall say that V is a *strongly stable* or a *rigid* submanifold of W if no small deformation of the complex structure of W changes the complex structure of a *neighborhood of* V *in* W. Alternatively we call V *a weakly stable* or simply a *stable* submanifold of W if no small deformation of the complex structure of W makes V disappear (see Definition 1 below). The rigidity of compact submanifolds of complex manifolds has been studied by A. Andreotti and E. Vesentini. The purpose of the present note is to discuss the (weak) stability of compact submanifolds of complex manifolds and some related topics.

1. Stability of compact submanifolds. In what follows we assume that all manifolds under consideration are paracompact and connected. By a *complex fibre manifold* we shall mean a complex manifold $\mathcal{W}$ together with a holomorphic map p of $\mathcal{W}$ onto a complex manifold B such that the rank of the Jacobian of p at each point of $\mathcal{W}$ is equal to the dimension of B. We call B the base space of the complex fibre manifold $\mathcal{W}$. Moreover, for any point $u \in B$, we denote the inverse image $p^{-1}(u)$ of u by W_u and call it the *fibre of* $\mathcal{W}$ *over* u. Obviously each fibre $W_u = p^{-1}(u)$ is a complex (analytic) submanifold of $\mathcal{W}$. We denote the complex fibre manifold $\mathcal{W}$ by the triple $(\mathcal{W}, B, p)$ when we want to indicate the map p and the base space B explicitly. Note that, in case the fibres W_u, $u \in B$, are compact, the triple $(\mathcal{W}, B, p)$ forms a complex analytic family of compact complex manifolds. For any subdomain N of B we call the complex fibre manifold $(p^{-1}(N), N, p)$ the *restriction* of $\mathcal{W}$ to N and denote it by the symbol $\mathcal{W}|N$. Let $\mathcal{V}$ be a complex submanifold of $\mathcal{W}|N$ such that $p(\mathcal{V}) = N$. We call $\mathcal{V}$ a *fibre submanifold* of the complex fibre manifold $\mathcal{W}|N$ if and only if the triple

* Received November 1, 1962.

[1] This research was partially supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under contract No. AF49(638)-253. Reproduction in whole or in part is permitted for any purpose of the United States Government.

$(\mathcal{V}, N, p)$ forms a complex fibre manifold. If, moreover, each fibre $V_u = \mathcal{V} \cap W_u$, $u \in N$, of $\mathcal{V}$ is compact, we call $\mathcal{V}$ a *fibre submanifold with compact fibres* of the complex fibre manifold $\mathcal{W} | N$.

Consider a compact complex submanifold V of a complex manifold W.

Definition 1. We call V a stable submanifold of W if and only if, for any complex fibre manifold $(\mathcal{W}, B, p)$ such that $p^{-1}(o) = W$ for a point $o \in B$, there exist a neighborhood N of o in B and a fibre submanifold $\mathcal{V}$ with compact fibres of the complex fibre manifold $\mathcal{W} | N$ such that $\mathcal{V} \cap W = V$.

In this section we shall prove the following theorem:

THEOREM 1. *Let Ψ be the sheaf over V of germs of holomorphic sections of the normal bundle of V in W. If the first cohomology group $H^1(V, \Psi)$ vanishes, then V is a stable submanifold of W.*

Let $(\mathcal{W}, B, p)$ be a complex fibre manifold such that $p^{-1}(o) = W$ for a point $o \in B$ and let $(u^1, u^2, \cdots, u^q)$ denote a local coordinate on B with the center o. Considering $V \subset W \subset \mathcal{W}$ as a submanifold of $\mathcal{W}$, we cover V by a finite number of coordinate neighborhoods $\mathcal{U}_i$ in $\mathcal{W}$ and choose a local coordinate

$$(z_i^1, \cdots, z_i^d, w_i^1, \cdots, w_i^r, u^1, \cdots, u^q) \tag{1}$$

on each neighborhood $\mathcal{U}_i$ such that the map p takes the form

$$p: (z_i^1, \cdots, z_i^d, w_i^1, \cdots, w_i^r, u^1, \cdots, u^q) \to (u^1, \cdots, u^q)$$

and such that the simultaneous equations

$$w_i^1 = \cdots = w_i^r = u^1 = \cdots = u^q = 0$$

define the submanifold V. For brevity we write

$$z_i = (z_i^1, \cdots, z_i^d), w_i = (w_i^1, \cdots, w_i^r), u = (u^1, \cdots, u^q),$$

and define

$$| z_i | = \max_\alpha | z_i^\alpha |, \ | w_i | = \max_\lambda | w_i^\lambda |, \ | u | = \max_\rho | u^\rho |.$$

We assume that each neighborhood $\mathcal{U}_i$ is a polycylinder consisting of all points (z_i, w_i, u), $| z_i | < 1$, $| w_i | < 1$, $| u | < 1$. On the intersection $\mathcal{U}_i \cap \mathcal{U}_k$ the local coordinates z_i^α, w_i^λ are holomorphic functions of z_k, w_k, u:

$$\begin{cases} z_i^\alpha = g_{ik}^\alpha(z_k, w_k, u), & \alpha = 1, 2, \cdots, d, \\ w_i^\lambda = f_{ik}^\lambda(z_k, w_k, u), & \lambda = 1, 2, \cdots, r. \end{cases} \tag{2}$$

Note that

$$f_{ik}^\lambda(z_k, 0, 0) = 0.$$

We set

$$U_k = V \cap \mathcal{U}_k.$$

We denote a point on V by z and, if $z = (z_k, 0, 0) \in U_k$, we call $z_k = (z_k^1, \cdots, z_k^d)$ the *local coordinate of z on U_k*. We define

$$a_{ik}{}^{\lambda}{}_{\mu}(z) = [\partial f_{ik}{}^{\lambda}(z_k, w_k, u)/\partial w_k{}^{\mu}]_{w_k=u=0},$$
$$b_{ik}{}^{\lambda}{}_{\rho}(z) = [\partial f_{ik}{}^{\lambda}(z_k, w_k, u)/\partial u^{\rho}]_{w_k=u=0},$$

and write

$$a_{ik}(z) = (a_{ik}{}^{\lambda}{}_{\mu}(z))_{\lambda,\mu=1,2,\cdots,r},$$
$$b_{ik}(z) = (b_{ik}{}^{\lambda}{}_{\rho}(z))_{\lambda=1,\cdots,r,\rho=1,\cdots,q}.$$

The *normal bundle* of V in $\mathcal{W}$ is defined by the system of transition matrices

$$\begin{pmatrix} a_{ik}(z) & b_{ik}(z) \\ 0 & 1_q \end{pmatrix}, \tag{3}$$

where 1_q denotes the $q \times q$ unit matrix. We denote by Φ the sheaf over V of germs of holomorphic sections of the normal bundle of V *in* $\mathcal{W}$. Ψ is, by definition, the sheaf over V of germs of holomorphic sections of the normal bundle of V *in* W which is determined by the system of transition matrices $a_{ik}(z)$. Hence we obtain the *exact sequence*

$$0 \longrightarrow \Psi \longrightarrow \Phi \xrightarrow{\kappa} \mathcal{O}^q \longrightarrow 0, \tag{4}$$

where $\mathcal{O}^q$ denotes the sheaf over V of germs of holomorphic sections of the trivial bundle $\boldsymbol{C}^q \times V$, $\boldsymbol{C}^q$ being the space of q complex variables. We assume that the first cohomology group $H^1(V, \Psi)$ vanishes. Hence we obtain from (4) the exact sequence

$$0 \longrightarrow H^0(V, \Psi) \longrightarrow H^0(V, \Phi) \xrightarrow{\kappa} \boldsymbol{C}^q \longrightarrow 0. \tag{5}$$

Note that any element ψ of $H^0(V, \Phi)$ is a collection $\{\psi_i(z)\}$ of vector-valued holomorphic functions

$$\psi_i(z) = (\psi_i^1(z), \cdots, \psi_i^r(z), \psi^{r+1}, \cdots, \psi^{r+q})$$

defined respectively on U_i satisfying

$$\psi_i^{\lambda}(z) = \sum_{\mu} a_{ik}{}^{\lambda}{}_{\mu}(z)\psi_k{}^{\mu}(z) + \sum_{\rho} b_{ik}{}^{\lambda}{}_{\rho}(z)\psi^{r+\rho}, \qquad \text{on } U_i \cap U_k,$$

where $\psi^{r+1}, \cdots, \psi^{r+q}$ are constants which are independent of i. Obviously we have

$$\kappa\psi = (\psi^{r+1}, \cdots, \psi^{r+q}). \tag{6}$$

In what follows we denote by t a point $(t_1, \cdots, t_n)$ in the space of

several complex variables $t_1, \cdots, t_n$ and set $|t| = \max_\nu |t_\nu|$. Moreover we denote by M_ϵ the polycylinder consisting of all points t, $|t| < \epsilon$, where ϵ is a *small* positive number. We denote by $T_t(M_\epsilon)$ the tangent space of M_ϵ at t. Consider an analytic family $\mathfrak{F}$ of compact complex submanifolds V_t, $t \in M_\epsilon$, of $\mathfrak{W}$ such that $V_0 = V$. $\mathfrak{F}$ is, by definition, a complex analytic submanifold of $\mathfrak{W} \times M_\epsilon$ such that $\mathfrak{F} \cap \mathfrak{W} \times t = V_t \times t$. We assume that $\mathfrak{F}$ is covered by the neighborhoods $\mathfrak{U}_i \times M_\epsilon$. Clearly the image $p(V_t)$ of each compact submanifold V_t is a single point on B. It follows that, on each neighborhood $\mathfrak{U}_i \times M_\epsilon$, the submanifold $\mathfrak{F}$ is defined by simultaneous holomorphic equations of the form

$$(7) \qquad \begin{cases} w_i^\lambda = \theta_i^\lambda(z_i, t), & \lambda = 1, 2, \cdots, r, \\ u^\rho = \theta^{r+\rho}(t), & \rho = 1, 2, \cdots, q. \end{cases}$$

For any tangent vector $\partial/\partial t = \sum c_\nu \partial/\partial t_\nu$ of M_ϵ at $t = 0$, we set

$$\begin{aligned} \psi_i^\lambda(z) &= [\partial\theta_i^\lambda(z_i, t)/\partial t]_{t=0}, & \lambda &= 1, 2, \cdots, r, \\ \psi^{r+\rho} &= [\partial\theta^{r+\rho}(t)/\partial t]_{t=0}, & \rho &= 1, \cdots, q \end{aligned}$$

and let

$$\psi_i(z) = (\psi_i^1(z), \cdots, \psi_i^r(z), \psi^{r+1}, \cdots, \psi^{r+q}).$$

The collection $\{\psi_i(z)\}$ of $\psi_i(z)$ represents an element of $H^0(V, \Phi)$. We define

$$(8) \qquad [\partial V_t/\partial t]_{t=0} = \{\psi_i(z)\}$$

and call it the *infinitesimal displacement* of V_t *along* $\partial/\partial t$ (see [2], Definition 4).

THEOREM 2. *There exists an analytic family $\mathfrak{F}$ of compact complex submanifolds V_t, $t \in M_\epsilon$, $\epsilon > 0$, of $\mathfrak{W}$ such that $V_0 = V$ and such that the map: $\partial/\partial t \to [\partial V_t/\partial t]_{t=0}$ maps the tangent space $T_0(M_\epsilon)$ isomorphically onto $H^0(V, \Phi)$ provided that the cohomology group $H^1(V, \Psi)$ vanishes.*

Proof. This theorem can be derived readily from Theorem 1* of [2]. Let $n = \dim H^0(V, \Phi)$. In view of (5) and (6) we can choose a base $\{\beta_1, \cdots, \beta_\nu, \cdots, \beta_n\}$ of the linear space $H^0(V, \Phi)$ such that

$$(9) \qquad \beta_\nu^{r+\rho} = \begin{cases} 1, & \text{if } \nu = n - q + \rho, \\ 0, & \text{otherwise.} \end{cases}$$

For brevity we set

$$\begin{aligned} w_i^{r+\rho} &= u^\rho, & \text{for } \rho &= 1, 2, \cdots, q, \\ w_i &= (w_i^1, \cdots, w_i^r, \cdots, w_i^{r+q}), \\ f_{ik}^{r+\rho}(z_k, w_k) &= w_k^{r+\rho}, \\ f_{ik} &= (f_{ik}^1, \cdots, f_{ik}^r, \cdots, f_{ik}^{r+q}), \\ g_{ik} &= (g_{ik}^1, \cdots, g_{ik}^d), \end{aligned}$$

and rewrite the equalities (2) in the form

$$\begin{cases} z_i = g_{ik}(z_k, w_k), \\ w_i = f_{ik}(z_k, w_k). \end{cases}$$

Our purpose is to construct on each neighborhood $\mathfrak{U}_i \times M_\epsilon$ a vector-valued holomorphic function of the form

$$(10) \qquad \phi_i(z_i, t) = (\theta_i^1(z_i, t), \cdots, \theta_i^r(z_i, t), t_{n-q+1}, \cdots, t_n)$$

satisfying the boundary conditions

$$\phi_i(z_i, 0) = 0,$$
$$[\partial\phi_i(z_i, t)/\partial t_\nu]_{t=0} = \beta_{\nu i}(z), \qquad \nu = 1, 2, \cdots, n,$$

such that

$$(11) \qquad \phi_i(g_{ik}(z_k, \phi_k(z_k, t)), t) = f_{ik}(z_k, \phi_k(z_k, t)).$$

We write the power series expansion of $\phi_i(z_i, t)$ in $t_1, \cdots, t_n$ in the form

$$\phi_i(z_i, t) = \phi_{i|1}(z_i, t) + \phi_{i|2}(z_i, t) + \cdots + \phi_{i|m}(z_i, t) + \cdots,$$

where each term $\phi_{i|m}(z_i, t)$ represents a homogeneous polynomial of degree m in $t_1, \cdots, t_n$, and we set

$$\phi_i^m(z_i, t) = \phi_{i|1}(z_i, t) + \phi_{i|2}(z_i, t) + \cdots + \phi_{i|m}(z_i, t).$$

Then the equality (11) is reduced to the system of congruences

$$(12)_m \qquad \phi_i^m(g_{ik}(z_k, \phi_k^m(z_k, t)), t) \equiv f_{ik}(z_k, \phi_k^m(z_k, t)) \pmod{t^{m+1}},$$
$$m = 1, 2, 3, \cdots.$$

Assuming that the polynomials $\phi_i^m(z_i, t)$ satisfying $(12)_m$ are obtained for an integer $m \geqq 1$ we determine homogeneous polynomials $\psi_{ik}(z, t)$ of degree $m+1$ in $t_1, \cdots, t_n$ by the congruences

$$\psi_{ik}(z, t) \equiv \phi_i^m(g_{ik}(z_k, \phi_k^m(z_k, t)), t) - f_{ik}(z_k, \phi_k^m(z_k, t)) \pmod{t^{m+2}}.$$

The collection $\{\psi_{ik}(z, t)\}$ of $\psi_{ik}(z, t)$ can be considered as a homogeneous polynomial of degree $m+1$ in $t_1, \cdots, t_n$ whose coefficients are 1-cocycles with coefficients in the sheaf Φ (see [2], § 2). Thus the collection $\{\psi_{ik}(z, t)\}$ determines a homogeneous polynomial $\psi_{m+1}(t)$ of degree $m+1$ in $t_1, \cdots, t_n$ with coefficients in $H^1(V, \Phi)$. The homogeneous polynomial $\psi_{m+1}(t)$ is called the m-th *obstruction.*

In order to construct the holomorphic functions $\phi_i(z_i, t)$ satisfying (11) it suffices to show that *the obstruction* $\psi_{m+1}(t)$ *vanishes for each integer*

$m \geqq 1$ (see [2], Theorem 1*). To show the vanishing of $\psi_{m+1}(t)$ we impose on $\phi_i^m(z_i, t)$ an additional restriction that

$$(13)_m \qquad \phi_i^m(z_i, t) = (\cdot \cdot \cdot, t_{n-q+1}, t_{n-q+2}, \cdot \cdot \cdot, t_n).$$

Then, since $f_{ik}^{r+\rho}(z_k, w_k) = w_k^{r+\rho}$, we obtain

$$\psi_{ik}^{r+\rho}(z, t) = 0, \qquad \text{for } \rho = 1, 2, \cdot \cdot \cdot, q.$$

Thus the coefficients of the homogeneous polynomial $\{\psi_{ik}(z, t)\}$ are 1-cocycles with coefficients in the *subsheaf* Ψ of Φ. By hypothesis the cohomology group $H^1(V, \Psi)$ vanishes. Hence the obstruction $\psi_{m+1}(t)$ vanishes. Moreover the 1-cocycle $\{\psi_{ik}(z, t)\}$ is a coboundary of a 0-cochain $\{\chi_i(z, t)\}$ composed of homogeneous polynomials $\chi_i(z, t)$ of degree $m + 1$ in $t_1, \cdot \cdot \cdot, t_n$ whose coefficients are sections of the sheaf Ψ over U_i. Thus

$$\chi_i(z, t) = (\chi_i^1(z, t), \cdot \cdot \cdot, \chi_i^r(z, t))$$

and

$$\psi_{ik}^\lambda(z, t) = \sum_{\mu=1}^{r} a_{ik}{}^\lambda{}_\mu(z) \chi_k^\mu(z, t) - \chi_i^\lambda(z, t), \qquad \text{for } \lambda = 1, \cdot \cdot \cdot, r.$$

We define

$$\phi_{i|m+1}(z_i, t) = (\chi_i^1(z, t), \cdot \cdot \cdot, \chi_i^r(z, t), 0, 0, \cdot \cdot \cdot, 0)$$

and set

$$\phi_i^{m+1}(z_i, t) = \phi_i^m(z_i, t) + \phi_{i|m+1}(z_i, t).$$

The polynomials $\phi_i^{m+1}(z_i, t)$ thus obtained satisfy the congruence $(12)_{m+1}$ and the condition $(13)_{m+1}$. We remark that, by (9), the linear terms

$$\phi_i^1(z_i, t) = \sum_{\nu=1}^{n} \beta_{\nu i}(z) t_\nu$$

satisfy the condition $(13)_1$. This completes our inductive proof of the vanishing of $\psi_{m+1}(t)$, q. e. d.

Remark. In view of (10) the family $\mathfrak{F}$ is defined on each neighborhood $\mathcal{U}_i \times M_\epsilon$ by the simultaneous equations

$$(14) \qquad \begin{cases} w_i^\lambda = \theta_i^\lambda(z_i, t), & \lambda = 1, 2, \cdot \cdot \cdot, r, \\ u^\rho = t_{n-q+\rho}, & \rho = 1, 2, \cdot \cdot \cdot, q. \end{cases}$$

Proof of Theorem 1. Let N_ϵ denote the neighborhood of o in B consisting of all points $u = (u^1, \cdot \cdot \cdot, u^q)$, $|u| < \epsilon$, where ϵ is a small positive number. Considering the family $\mathfrak{F}$ of compact complex submanifolds V_t, $t \in M_\epsilon$, of $\mathcal{W}$ defined by (14), we define a linear map: $u \rightarrow t(u)$ of N_ϵ into M_ϵ by setting

$$(15) \qquad t(u) = (0, 0, \cdot \cdot \cdot, 0, u^1, u^2, \cdot \cdot \cdot, u^q).$$

Then the union $\mathfrak{V} = \bigcup_u V_{t(u)}$ of the submanifolds $V_{t(u)}$, $u \in N_\epsilon$, of $\mathfrak{W}$ forms a fibre submanifold with compact fibres of the complex fibre manifold $\mathfrak{W} | N_\epsilon$ and $\mathfrak{V} \cap W = V$. Thus we conclude that V is a stable submanifold of W, q. e. d.

Let $(\mathfrak{W}, B, p)$ be a complex fibre manifold and let $W = p^{-1}(o)$ be the fibre of $\mathfrak{W}$ over a point $o \in B$. Moreover, let V be a compact complex submanifold of W and denote by Ψ the sheaf of germs of holomorphic sections of the normal bundle of V in W. The above Theorem 1 asserts that, if $H^1(V, \Psi)$ vanishes, then, for a sufficiently small neighborhood N of o in B, there exists a fibre submanifold $\mathfrak{V}$ with compact fibres of the complex fibre manifold $\mathfrak{W} | N$ such that $\mathfrak{V} \cap W = V$. The fibre submanifold $\mathfrak{V}$ is, in general, not unique. In fact, in the above construction of $\mathfrak{V} = \bigcup_u V_{t(u)}$, we may replace (15) by

$$(15)' \qquad t(u) = (t^1(u), \cdots, t^\lambda(u), \cdots, t^r(u), u^1, \cdots, u^q)$$

where the $t^\lambda(u)$ are holomorphic functions in u such that $t^\lambda(0) = 0$, $| t^\lambda(u) | < \epsilon$ for $| u | < \epsilon$.

THEOREM 3. *If the cohomology groups $H^1(V, \Psi)$ and $H^0(V, \Psi)$ both vanish, then, for a sufficiently small neighborhood N of o in B, there exists one and only one fibre submanifold $\mathfrak{V}$ with compact fibres of the complex fibre manifold $\mathfrak{W} | N$ such that $\mathfrak{V} \cap W = V$.*

Proof. It suffices to prove the uniqueness of $\mathfrak{V}$. In terms of the local coordinate (1) the submanifold $\mathfrak{V}$ is defined on each neighborhood $\mathfrak{U}_i$ by simultaneous equations of the form

$$w_i^\lambda = \theta_i^\lambda(z_i, u), \qquad \lambda = 1, 2, \cdots, r.$$

Writing t for u we consider $\mathfrak{V}$ as an analytic family consisting of compact submanifolds $V_t = \mathfrak{V} \cap p^{-1}(t)$, $t \in N$ of $\mathfrak{W}$. Obviously the submanifold V_t is defined on each neighborhood $\mathfrak{U}_i$ by the simultaneous equations

$$\begin{cases} w_i^\lambda = \theta_i^\lambda(z_i, t), & \lambda = 1, 2, \cdots, r, \\ u^\rho = t^\rho, & \rho = 1, 2, \cdots, q. \end{cases}$$

We compute the infinitesimal displacement $[\partial V_t / \partial t]_{t=0}$ and find that

$$(16) \qquad \kappa [\partial V_t / \partial t]_{t=0} = (c_1, c_2, \cdots, c_q), \qquad \text{for } \partial/\partial t = \sum c_\rho \partial / \partial t^\rho.$$

Since, by hypothesis, $H^0(V, \Psi)$ vanishes, we infer from (5) that the map

$$\kappa \colon H^0(V, \Phi) \to \boldsymbol{C}^q$$

is bijective. Combining this with (16) we see that the map: $\partial/\partial t \to [\partial V_t/\partial t]_{t=0}$ maps the tangent space $T_0(N)$ of N isomorphically onto $H^0(V, \Phi)$. Consequently the *analytic family* $\mathfrak{V}$ *of the compact submanifolds* V_t, $t \in N$, *of* $\mathfrak{W}$ *is maximal at* $t=0$ (see [2], Theorem 2). Now let $\mathfrak{V}'$ be another fibre submanifold with compact fibres of $\mathfrak{W}|N$ such that $\mathfrak{V}' \cap W = V$ and let $V_t' = \mathfrak{V}' \cap p^{-1}(t)$ for $t \in N$. Since the family $\mathfrak{V}$ is maximal at $t=0$, there exists a holomorphic map $h: t \to h(t)$ of N into N such that $V_t' = V_{h(t)}$. It follows that $t = p(V_t') = p(V_{h(t)}) = h(t)$ and therefore $V_t' = V_t$. Thus we see that $\mathfrak{V}'$ coincides with $\mathfrak{V}$, q.e.d.

EXAMPLE. *Non-singular rational curves on complex analytic surfaces.* Let C be a non-singular rational curve on a complex analytic surface S and let $i(C, C)$ denote the intersection multiplicity of C with itself. Letting Ψ be the sheaf over C of germs of holomorphic sections of the normal bundle of C in S, we infer readily that $H^1(C, \Psi)$ vanishes if and only if $i(C, C) \geqq -1$. Hence we conclude that, *if* $i(C, C) \geqq -1$, C *is a stable submanifold of* S.

A simple example of surface containing a non-singular rational curve C with $i(C, C) \leqq 0$ is given by a rational ruled surface. Let P be a projective line and denote by ζ a non-homogeneous coordinate on P. Moreover let U_1 and U_2 be two copies of the space $\boldsymbol{C}$ of a complex variable z. For each integer $n \geqq 0$ we form a rational ruled surface

$$S_n = P \times U_1 \cup P \times U_2$$

by identifying $(\zeta_1, z_1) \in P \times U_1$ with $(\zeta_2, z_2) \in P \times U_2$ if and only if

$$\begin{cases} \zeta_1 = z_2{}^n \zeta_2, \\ z_1 = 1/z_2, \end{cases}$$

and we denote by C the curve on S_n defined by the equations $\zeta_1 = \zeta_2 = 0$. Clearly C is a *non-singular rational curve with* $i(C, C) = -n$. Moreover, for any irreducible curve Γ on S_n, we have the *inequality*: $i(\Gamma, \Gamma) \geqq -n$.

Now we show that, *in case* $n \geqq 2$, C is an *unstable submanifold of* S_n. For each complex number t we define a surface

$$S_{n,t} = P \times U_1 \cup P \times U_2$$

by identifying $(\zeta_1, z_1) \in P \times U_1$ with $(\zeta_2, z_2) \in P \times U_2$ if and only if

$$\begin{cases} \zeta_1 = z_2{}^n \zeta_2 + t z_2{}^k, \\ z_1 = 1/z_2, \end{cases} \tag{17}$$

where k is a fixed positive integer $\leqq \frac{1}{2}n$. The surface $S_{n,t}$ depends holomorphically on t. Thus the set of all surfaces $S_{n,t}$ form an analytic family.

It is obvious that $S_{n,0}$ coincides with S_n, while, for $t \neq 0$, $S_{n,t}$ is complex analytically homeomorphic to the ruled surface S_{n-2k}. In fact, introducing new variables

$$\zeta_1' = (z_1^k \zeta_1 - t)/t\zeta_1,$$
$$\zeta_2' = \zeta_2/(tz_2^{n-k}\zeta_2 + t^2),$$

we infer readily that the equalities (17) are equivalent to

$$\begin{cases} \zeta_1' = z_2^{n-2k}\zeta_2', \\ z_1 = 1/z_2. \end{cases}$$

It follows that there exists on $S_{n,t}$, $t \neq 0$, no irreducible curve C_t with $i(C_t, C_t) = -n$. Thus we see that the curve C on S_n is unstable.

2. Stability of fibre structures. Let W be a complex manifold. By a *fibre structure* on W we shall mean a pair (B, p) of a complex manifold B and a holomorphic map p of W onto B such that the triple (W, B, p) forms a complex fibre manifold. A (complex) analytic family of compact complex manifolds is, by definition, a complex fibre manifold with compact fibres. Let $\mathcal{W} = (\mathcal{W}, S, \varpi)$ be an analytic family of compact complex manifolds. By an *analytic family of fibre structures on the analytic family* $\mathcal{W}$ we shall mean a pair $(\mathcal{B}, \mathcal{P})$ of an analytic family $\mathcal{B} = (\mathcal{B}, \mathcal{S}, \pi)$ of compact complex manifolds and a holomorphic map $\mathcal{P}$ of $\mathcal{W}$ onto $\mathcal{B}$ such that $\pi\mathcal{P} = \varpi$ and such that the triple $(\mathcal{W}, \mathcal{B}, \mathcal{P})$ forms a complex fibre manifold. For each point s of S we set $W_s = \varpi^{-1}(s)$, $B_s = \pi^{-1}(s)$ and denote by p_s the restriction of the map $\mathcal{P}$ to the submanifold W_s of $\mathcal{W}$. Obviously the pair (B_s, p_s) defines a fibre structure on W_s. We call (B_s, p_s) the *fibre* over s of the family $(\mathcal{B}, \mathcal{P})$.

Definition 2. A fibre structure (B, p) on a compact complex manifold W is said to be stable if and only if, for any analytic family $\mathcal{W} = (\mathcal{W}, S, \varpi)$ of compact complex manifolds such that $\varpi^{-1}(o) = W$ for a point $o \in S$, there exist a neighborhood N of o in S and an analytic family $(\mathcal{B}, \mathcal{P})$ of fibre structures on the analytic family $\mathcal{W}|N$ of which the fibre (B_o, p_o) over o coincides with (B, p).

A compact complex manifold V is said to be *regular* if and only if the first cohomology group $H^1(V, \mathcal{O})$ of V with coefficients in the sheaf $\mathcal{O}$ over V of germs of holomorphic functions vanishes.

THEOREM 4. *Let $(W, B\ p)$ be a compact complex fibre manifold. If all fibres $p^{-1}(b)$, $b \in B$, are regular, then the fibre structure (B, p) on W is stable.*

Proof. Let $(\mathcal{W}, S, \varpi)$ be an analytic family of compact complex manifolds such that $\varpi^{-1}(o) = W$ for a point $o \in S$ and let $(u^1, u^2, \cdots, u^q)$ denote a local coordinate on S with the center o. We take a point $b \in B$ and let $V = p^{-1}(b)$. We cover V by a finite number of coordinate neighborhoods $\mathcal{U}_i$ *in* $\mathcal{W}$ and choose a local coordinate

$$(z_i^1, \cdots, z_i^d, w_i^1, \cdots, w_i^r, u^1, \cdots, u^q)$$

on each neighborhood $\mathcal{U}_i$ such that the map ϖ takes the form

$$\varpi : (z_i^1, \cdots z_i^d, w_i^1, \cdots, w_i^r, u^1, \cdots, u^q) \to (u^1, \cdots, u^q)$$

and such that the simultaneous equations

$$w_i^1 = \cdots = w_i^r = u^1 = \cdots = u^q = 0$$

define the submanifold V. Since V is a fibre of the fibre manifold W, we may assume further that

$$w_i^\lambda = w_k^\lambda, \qquad \text{on } W \cap \mathcal{U}_i \cap \mathcal{U}_k.$$

For brevity we write $z_i = (z_i^1, \cdots, z_i^d)$, $w_i = (w_i^1, \cdots, w_i^r)$ and $u = (u^1, \cdots, u^q)$. The normal bundle of V in $\mathcal{W}$ is defined by a system of transition matrices of the form

$$\begin{pmatrix} 1_r & b_{ik} \\ 0 & 1_q \end{pmatrix}$$

(compare (3)). Let Φ be the sheaf over V of germs of holomorphic sections of the normal bundle of V in $\mathcal{W}$ and let $\mathcal{O}^\nu$ denote the sheaf over V of germs of holomorphic sections of the trivial bundle $C^\nu \times V$. Then we have the exact sequence

$$0 \longrightarrow \mathcal{O}^r \longrightarrow \Phi \xrightarrow{\kappa} \mathcal{O}^q \longrightarrow 0. \tag{18}$$

Since, by hypothesis, the fibre V is regular, the cohomology group $H^1(V, \mathcal{O}^\nu)$ vanishes for $\nu = 1, 2, 3, \cdots$. Hence we infer from (18) that the sequence

$$0 \longrightarrow C^r \longrightarrow H^0(V, \Phi) \longrightarrow C^q \longrightarrow 0 \tag{19}$$

is exact and that $H^1(V, \Phi)$ vanishes. By virtue of a theorem of completeness of characteristic systems (see [2] § 1), we conclude therefore the existence of an analytic family $\mathcal{F}$ of compact complex submanifolds V_t, $t \in M_\epsilon$, of $\mathcal{W}$ such that $V_0 = V$ and such that the map: $\partial/\partial t \to [\partial V_t/\partial t]_{t=0}$ maps the tangent space $T_0(M_\epsilon)$ isomorphically onto $H^0(V, \Phi)$, where M_ϵ is the polycylinder

of center 0 and of radius $\epsilon > 0$ in the space of $r+q$ complex variables $t_1, \cdots, t_r, t_{r+1}, \cdots, t_{r+q}$.

Now we examine the structure of $\mathfrak{F}$. On each neighborhood $\mathfrak{U}_i$ the submanifold V_t is defined by simultaneous holomorphic equations of the form

$$\left\{\begin{array}{ll} w_i^\lambda = \theta_i^\lambda(z_i, t), & \lambda = 1, 2, \cdots, r, \\ u^\rho = \theta^{r+\rho}(t), & \rho = 1, 2, \cdots, q, \end{array}\right. \tag{20}$$

and

$$[\partial V_t/\partial t]_{t=0} = (\psi_i^1(z), \cdots, \psi_i^r(z), \psi^{r+1}, \cdots, \psi^{r+q}),$$

where

$$\psi_i^\lambda(z) = [\partial\theta_i^\lambda(z_i, t)/\partial t]_{t=0}, \qquad \psi^{r+\rho} = [\partial\theta^{r+\rho}(t)/\partial t]_{t=0}.$$

In view of the exact sequence (19), we may assume therefore that

$$\begin{array}{ll} [\partial\theta_i^\lambda(z_i, t)/\partial t_\nu]_{t=0} = \delta_\nu^\lambda, & \text{for } \nu = 1, 2, \cdots, r, \\ [\partial\theta^{r+\rho}(t)/\partial t_\nu]_{t=0} = \delta_\nu^{r+\rho}, & \text{for } \nu = 1, \cdots, r, \cdots, r+q. \end{array}$$

These equalities show that *the simultaneous equations* (20) *are solvable with respect to the variables* $t_1, t_2, \cdots, t_{r+q}$; namely (20) are equivalent to the simultaneous holomorphic equations of the form

$$t_\nu = \tau_\nu(z_i, w_i, u), \qquad \nu = 1, 2, \cdots, r+q. \tag{21}$$

Thus we conclude that *the union* $\mathfrak{U} = \bigcup_t V_t$ *of the submanifolds* V_t, $t \in M_\epsilon$, *forms a neighborhood of* V *in* $\mathfrak{W}$ *and that there is a holomorphic map* τ *of* $\mathfrak{U}$ *onto* M_ϵ *which maps* V_t *onto* t, *provided that the positive number* ϵ *is sufficiently small.*

We construct for each point $b \in B$ a neighborhood $\mathfrak{U}(b)$ of $p^{-1}(b)$, a polycylinder $M(b)$ and a holomorphic map τ_b of $\mathfrak{U}(b)$ onto $M(b)$ in the manner described above. Moreover we take in each polycylinder $M(b)$ a sufficiently small concentric polycylinder $C(b)$ and let $\mathfrak{V}(b) = \tau_b^{-1}C(b)$. Obviously $\mathfrak{V}(b)$ is a neighborhood of $p^{-1}(b)$. Now we select a finite number of neighborhoods $\mathfrak{V}(b_j)$, $j = 1, 2, \cdots, l$, which cover W in such a way that, if $\mathfrak{V}(b_j) \cap \mathfrak{V}(b_k)$ is non-empty, then $\mathfrak{V}(b_j) \subset \mathfrak{U}(b_k)$ and $\mathfrak{V}(b_k) \subset \mathfrak{U}(b_j)$. For brevity we write $\mathfrak{V}_j = \mathfrak{V}(b_j)$, $C_j = C(b_j)$ and denote by τ_j the restriction of τ_{b_j} to $\mathfrak{V}_j$. Moreover we set

$$C_{jk} = \tau_j(\mathfrak{V}_j \cap \mathfrak{V}_k),$$

provided that $\mathfrak{V}_j \cap \mathfrak{V}_k$ is non-empty. Consider a point t in $C_{jk} \subset C_j$. Since $\mathfrak{V}_j \subset \mathfrak{U}(b_k)$, $\tau_j^{-1}(t)$ is a compact complex submanifold of $\mathfrak{U}(b_k)$, while τ_{b_k} is a holomorphic map of $\mathfrak{U}(b_k)$ onto $M(b_k)$. Hence $\tau_{b_k}\tau_j^{-1}(t)$ is a single point

on $M(b_k)$ which is obviously contained in the subdomain C_k of $M(b_k)$. Thus we infer that $\tau_k \tau_j^{-1}$ is a biholomorphic map of $C_{jk} \subset C_j$ onto $C_{kj} \subset C_k$. Consequently, identifying $C_{jk} \subset C_j$ with $C_{kj} \subset C_k$ by means of the map $\tau_k \tau_j^{-1}$, we obtain from the collection of the polycylinders C_j a complex manifold $C = \bigcup_j C_j$. We form the union $\mathfrak{V} = \bigcup_j \mathfrak{V}_j$ of $\mathfrak{V}_j$. Obviously $\mathfrak{V}$ is an open subset of $\mathfrak{W}$ which contains W. We define a holomorphic map $\mathfrak{P}$ of $\mathfrak{V}$ onto C by setting $\mathfrak{P} = \tau_j$ on each $\mathfrak{V}_j$. Let

$$\pi = \varpi \mathfrak{P}^{-1}.$$

It is clear by (20) and (21) that π is a holomorphic map of C into S.

Let N be a neighborhood of o in S such that $\varpi^{-1}(N) \subset \mathfrak{V}$ and let $\mathfrak{B} = \pi^{-1}(N) = \mathfrak{P}\varpi^{-1}(N)$. We infer readily that the triple $(\mathfrak{B}, N, \pi)$ forms an analytic family of compact complex manifolds and that the pair $(\mathfrak{B}, \mathfrak{P})$ defines an analytic family of fibre structures on the analytic family $\mathfrak{W} \,|\, N = (\varpi^{-1}(N), N, \varpi)$. It is clear that the fibre (B_o, p_o) of $(\mathfrak{B}, \mathfrak{P})$ over o coincides with the fibre structure (B, p), q. e. d.

Remark. For any compact complex manifold V which is *irregular* in the sense that $H^1(V, \mathfrak{O}) \neq 0$, $\mathfrak{O}$ being the sheaf over V of germs of holomorphic functions, there exists a compact complex manifold W with an *unstable* fibre structure (B, p) such that the fibres $p^{-1}(b)$, $b \in B$, are complex analytically homeomorphic to V. In fact, an example of such a manifold is given by the product $W = P \times V$ of V and a projective line P with the fibre structure defined by the canonical projection p of W onto P. In order to show that the fibre structure (P, p) of W is unstable, we construct an analytic family of projective line bundles over V in the following manner: Let $\{U_i\}$ be a finite covering of V by polycylinders U_i and let $\{h_{ik}\}$ be a 1-cocycle composed of holomorphic functions $h_{ik} = h_{ik}(z)$ on $U_i \cap U_k$ which represents an element $h \neq 0$ of $H^1(V, \mathfrak{O})$. For each complex number t we define a compact complex manifold

$$W_t = \bigcup_i P \times U_i$$

by identifying $(\zeta_i, z) \in P \times U_i$ with $(\zeta_k, z) \in P \times U_k$ if and only if

$$\zeta_i = \zeta_k \cdot \exp t h_{ik}(z).$$

It is clear that $(\zeta_i, z) \to z$ defines a holomorphic map of W_t onto V and thus W_t is a projective line bundle over V. Obviously the set of all projective line bundles W_t, $t \in \boldsymbol{C}$, forms an analytic family and W_0 coincides with W. The projective line bundle W_t is trivial if and only if t belongs to a discontinuous subgroup Δ of the additive group $\boldsymbol{C}$. Thus, for $t \notin \Delta$, W_t has no holomorphic

section: $z \to (\zeta_i(z), z)$ with $\zeta_i(z) \neq 0$, $\neq \infty$. This shows that each fibre $p^{-1}(\zeta) = \zeta \times V$, $\zeta \neq 0$, $\neq \infty$, of the complex fibre manifold (W, P, p) is unstable. It follows immediately that the fibre structure (P, p) of W is unstable.

3. Exceptional submanifolds. In what follows we denote by P_n a complex projective n-space, $n \geqq 1$. By a P_n-bundle we mean a complex analytic fibre bundle of which the fibre is P_n and the structure group is a projective transformation group acting on P_n. We note that any complex fibre manifold whose fibres are complex analytically homeomorphic to P_n is a P_n-bundle (see [3], p. 448). We denote by $E(P_n)$ the fundamental complex line bundle on P_n, i. e., the complex line bundle determined by a hyperplane in P_n. Let V be a compact complex submanifold of codimension 1 of a complex manifold W and let F denote the normal bundle of V in W. Note that F is a complex line bundle over V.

Definition 3. We call V an exceptional submanifold of W if and only if V satisfies the following conditions: i) V is the bundle space of a P_n-bundle; thus there exists a holomorphic map p of V onto a compact complex manifold B such that, for each point $b \in B$, $p^{-1}(b)$ is (complex analytically homeomorphic to) P_n; ii) the restriction $F \mid p^{-1}(b)$ of F to each fibre $p^{-1}(b)$ of V is equal to the inverse of the fundamental complex line bundle $E(P_n)$ on the fibre $p^{-1}(b) = P_n$.

THEOREM 5. *Every exceptional submanifold of a complex manifold is stable. Let $(\mathcal{W}, S, \varpi)$ be a complex fibre manifold and let V be an exceptional submanifold of a fibre $W = \varpi^{-1}(o)$, $o \in S$, of $\mathcal{W}$. Then, for a sufficiently small neighborhood N of o in S, there exists one and only one fibre submanifold $\mathcal{V}$ of the complex fibre manifold $\mathcal{W} \mid N$ such that $\mathcal{V} \cap W = V$. Moreover each fibre $V_s = \mathcal{V} \cap \varpi^{-1}(s)$, $s \in N$, of $\mathcal{V}$ is an exceptional submanifold of $\varpi^{-1}(s)$.*

Proof. Let F be the normal bundle of V in W and let Ψ be the sheaf over V of germs of holomorphic sections of F. By hypothesis, there is a holomorphic map p of V onto a compact complex manifold B such that $p^{-1}(b) = P_n$ and $F \mid p^{-1}(b) = -E(P_n)$ for each point $b \in B$. It follows that the cohomology group $H^q(V, \Psi)$ vanishes for $q = 0, 1, 2, \cdots$. Hence, by Theorem 1, V is a stable submanifold of W. Moreover, by Theorem 3, there exists one and only one fibre submanifold $\mathcal{V}$ of $\mathcal{W} \mid N$ such that $\mathcal{V} \cap W = V$.

Now, by Theorem 4, the fibre structure (B,p) on V is stable. Thus there exists an analytic family $(\mathfrak{B},\mathfrak{P})$ of fibre structures on the analytic family $\mathfrak{V}$ of which the fibre (B_o,p_o) over o coincides with (B,p). Consider the fibre (B_s,p_s) of $(\mathfrak{B},\mathfrak{P})$ over $s\in N$. Every fibre $p_o^{-1}(b)$, $b\in B_o$, of $V=\mathfrak{V}\cap\varpi^{-1}(o)$ is complex analytically homeomorphic to P_n, while the complex structure of P_n is rigid (see Frölicher and Nijenhuis [1]). Hence we conclude that every fibre $p_s^{-1}(b)$, $b\in B_s$, of each fibre manifold $V_s=\mathfrak{V}\cap\varpi^{-1}(s)$, $s\in N$, is complex analytically homeomorphic to P_n. Thus the fibre manifold V_s is a P_n-bundle. It follows immediately that V_s is an exceptional submanifold of $\varpi^{-1}(s)$, q. e. d.

4. Monoidal transforms. Let Z be a compact complex manifold. For any compact complex submanifold C of Z of codimension $n\geqq 2$ we denote by μ_C the *monoidal transformation with the center* C. The *monoidal transform* $\mu_C(Z)$ of Z is a compact complex manifold and the *total transform* $\mu_C(C)$ of C is an execption submanifold of $\mu_C(Z)$. Moreover the inverse μ_C^{-1} of μ_C is a holomorphic map of $\mu_C(Z)$ onto Z which maps $\mu_C(Z)-\mu_C(C)$ biholomorphically onto $Z-C$.

Consider an analytic family $\mathfrak{C}$ of compact complex submanifolds C_s, $s\in N$, of Z, where N denotes a connected complex manifold. $\mathfrak{C}$ is, by definition, a fibre submanifold of the complex fibre manifold $(Z\times N,N,\pi)$, where π denotes the canonical projection of $Z\times N$ onto N, and

$$C_s\times s=\mathfrak{C}\cap Z\times s, \qquad \text{for } s\in N.$$

Let $\mu(Z\times N)$ denote the monoidal transform of $Z\times N$ with respect to the center $\mathfrak{C}$. The fibre structure (N,π) on $Z\times N$ induces a fibre structure (N,Π) on $\mu(Z\times N)$ such that

$$\Pi^{-1}(s)=\mu_{C_s}(Z), \qquad \text{for } s\in N.$$

Thus the triple $(\mu(Z\times N),N,\Pi)$ forms an analytic family consisting of the monoidal transforms $\mu_{C_s}(Z)$, $s\in N$, of Z.

Definition 4. The monoidal transform $\mu_C(Z)$ of Z will be said to be stable if and only if, for any analytic family $(\mathfrak{W},S,\varpi)$ of compact complex manifolds such that $\varpi^{-1}(o)=\mu_C(Z)$ for a point $o\in S$, there exist a neighborhood N of o in S and an analytic family $\mathfrak{C}$ of compact complex submanifolds C_s, $s\in N$, of Z such that $C_o=C$ and such that the restriction $(\mathfrak{W}\,|\,N,N,\varpi)$ of $(\mathfrak{W},S,\varpi)$ coincides with the analytic family $(\mu(Z\times N),N,\Pi)$ of the monoidal transforms $\mu_{C_s}(Z)$, $s\in N$, of Z.

THEOREM 6. *Let Θ denote the sheaf over Z of germs of holomorphic vector fields. If the first cohomology group $H^1(Z,\Theta)$ vanishes then any monoidal transform $\mu_C(Z)$ of Z is stable.*

Proof. We set $W=\mu_C(Z)$, $f=\mu_C^{-1}$ and let G denote the graph of the holomorphic map f of W onto Z. The graph G is, by definition, the submanifold of $W\times Z$ consisting of all points $(w,f(w))$, $w\in W$. First we shall show that G is a stable submanifold of $W\times Z$. Let p denote the restriction to G of the canonical projection $W\times Z\to Z$. Obviously

$$h\colon w\to(w,f(w))$$

is a biholomorphic map of W onto G and

$$ph=f. \tag{22}$$

We fix a finite covering $\mathfrak{U}=\{U_i\}$ of Z by coordinate neighborhoods U_i with respective local coordinates $z_i=(z_i^1,\cdots,z_i^n,\cdots,z_i^d)$ such that each non-empty intersection $C\cap U_i$ coincides with the coordinate plane: $z_i^1=z_i^2=\cdots=z_i^n=0$. Moreover we assume that each U_i is the polycylinder $:|z_i|<1$. The equality (22) shows that p^{-1} is a monoidal transformation with the center C. Hence we conclude by an elementary consideration that

$$H^1(p^{-1}(U_i),\mathcal{O})=0, \tag{23}$$

where $\mathcal{O}$ denotes the sheaf over $p^{-1}(U_i)$ of germs of holomorphic functions.

Let Ψ denote the sheaf of germs of holomorphic sections of the normal bundle of G in $W\times Z$. The normal bundle of G in $W\times Z$ is induced from the tangent bundle of Z by the map p of G onto Z, while the tangent bundle of Z is trivial on each neighborhood U_i. Hence, by (23), the cohomology group $H^1(p^{-1}(U_i),\Psi)$ vanishes. Letting $\mathfrak{U}^*$ denote the covering $\{p^{-1}(U_i)\}$ of G, we infer therefore that

$$H^1(G,\Psi)\cong H^1(\mathfrak{U}^*,\Psi)\cong H^1(\mathfrak{U},\Theta)\cong H^1(Z,\Theta),$$

while, by hypothesis, $H^1(Z,\Theta)$ vanishes. Hence the cohomology group $H^1(G,\Psi)$ vanishes and consequently, by Theorem 1, G is a stable submanifold of $W\times Z$.

Let $(\mathcal{W},S,\varpi)$ be a complex analytic family of compact complex manifolds $W_s=\varpi^{-1}(s)$, $s\in S$, such that $W_o=W=\mu_C(Z)$ for a point $o\in S$. The total transform $\mu_C(C)$ is an exceptional submanifold of W_o. Hence, by Theorem 5, there exist a neighborhood N of o in S and a fibre submanifold $\mathcal{V}$ of the complex fibre manifold $\mathcal{W}|N$ such that $\mathcal{V}\cap W_o=\mu_C(C)$ and such

that each fibre $V_s = \mathcal{V} \cap W_s$, $s \in N$, is an exceptional submanifold of W_s. Thus V_s is (the bundle space of) a P_{n-1}-bundle over a compact complex manifold C_s. Moreover the manifolds C_s, $s \in N$, form a complex analytic family $(\mathcal{C}, N, \pi)$ and there exists a holomorphic map $\mathcal{P}$ of $\mathcal{V}$ onto $\mathcal{C}$ which reduces on each fibre V_s of $\mathcal{V}$ to the canonical projection of V_s onto C_s. Thus $\mathcal{V}$ is (the bundle space of) a P_{n-1}-bundle over $\mathcal{C}$. In what follows we always assume that N is *sufficiently small* and write for the sake of simplicity $\mathcal{W}$ in place of $\mathcal{W} | N$.

We consider $\mathcal{W} \times Z$ as a complex analytic family consisting of the product manifolds $W_s \times Z$, $s \in N$. Since G is a stable submanifold of $W_0 \times Z$, there exists a fibre submanifold $\mathcal{G}$ of the complex manifold $\mathcal{W} \times Z$ such that $\mathcal{G} \cap W_0 \times Z = G$. The submanifold $\mathcal{G}$ defines a holomorphic map F of $\mathcal{W}$ onto Z of which the graph is $\mathcal{G}$. Clearly F is an extension of the map f of W_0 onto Z which reduces on V_0 to the canonical projection of the P_{n-1}-bundle V_0 onto its base space $C_0 = C \subset Z$. Hence we infer that F maps each fibre of the P_{n-1}-bundle $\mathcal{V}$ onto a point in Z and that $F\mathcal{P}^{-1}$ is a holomorphic map of $\mathcal{C}$ into Z which is biholomorphic on each fibre C_s, $s \in N$, of $\mathcal{C}$. We identify C_s with the submanifold $F\mathcal{P}^{-1}(C_s) \times s$ of $Z \times N$ and consider $\mathcal{C}$ as an analytic family of compact submanifolds of Z. Moreover we define Π to be the holomorphic map

$$w \rightarrow \Pi(w) = (F(w), \varpi(w))$$

of $\mathcal{W}$ onto $Z \times N$ which maps each fibre W_s of $\mathcal{W}$ onto $Z \times s$. Clearly Π is an extension of the map $f = \mu_C^{-1}$ of W_0 onto $Z \times o = Z$ while, on the submanifold $\mathcal{V}$ of $\mathcal{W}$, the map Π coincides with the map $\mathcal{P}$ of $\mathcal{V}$ onto $\mathcal{C}$. Hence we conclude that Π^{-1} is a monoidal transformation with the center $\mathcal{C}$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. Frölicher and A. Nijenhuis, "A theorem on stability of complex structures," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 43 (1957), pp. 239-241.

[2] K. Kodaira, "A theorem of completeness of characteristic systems for analytic families of compact submanifolds of complex manifolds," *Annals of Mathematics*, vol. 75 (1962), pp. 146-162.

[3] K. Kodaira and D. C. Spencer, "On deformations of complex analytic ctructures," I and II, *Annals of Mathematics*, vol. 67 (1958), pp. 328-466.

ON A QUESTION OF BISHOP AND PHELPS.*

By Victor Klee.[1]

In a remarkable paper, Bishop and Phelps [1] settle several important questions concerning supporting hyperplanes of convex sets in Banach spaces. In particular, they show that if C is a closed convex subset of a Banach space E and $\emptyset \neq C \neq E$, then C has many support points. (These are points p of C which lie on supporting hyperplanes of C, or equivalently for which $fp = \inf fC$ for some $f \in E^* \sim \{0\}$.) Completeness of E is essential, for the author [3] has described a precompact closed convex set A in an incomplete inner product space such that A has no support points. (See also p. 158 of Bourbaki [2].) As is mentioned in [1] and detailed in [5], some of the machinery of [1] can be developed in an arbitrary complete locally convex space. However, this does not extend to the theorem quoted above, and both [1] and [5] ask whether a closed convex subset of a complete locally convex space must have support points. In the present paper, this question is answered negatively by means of an example in the cartesian space $R^{\aleph_0}$ (with its usual product topology). However, some interesting unsolved problems remain and these are mentioned at the end of the paper.

The points of the space $R^{\aleph_0}$ are infinite sequences $x = (x_1, x_2, \cdots)$ of real numbers. For each $n \in N$ (natural numbers), let L_n denote the n-dimensional linear subspace of $R^{\aleph_0}$ consisting of all points $x \in R^{\aleph_0}$ such that $x_i = 0$ for all $i > n$. Let π_n denote the natural projection of $R^{\aleph_0}$ onto L_n, so that $\pi_n(x) = (x_1, \cdots, x_n, 0, 0, \cdots) \in L_n$. We shall construct a closed convex proper subset C of $R^{\aleph_0}$ such that C contains the positive cone $P = \{x \in R^{\aleph_0} : x_i \geqq 0 \text{ for all } i\}$, and such that for each $n \in N$ the set $\pi_n C$ is an *open* subset of L_n. If a linear functional f on $R^{\aleph_0}$ is bounded below on C it is evident that $\inf fP = 0$, whence $f = g\pi_n$ for some $n \in N$ and $g \in L_n^*$. (For a proof of this, see Mackey [4].) Except when g is identically zero, the openness of $\pi_n C$ in L_n implies the openness of $g\pi_n C$ in R and thus f cannot attain a minimum on C. Since this is true even without assuming the continuity of the linear functional f, it follows in a very strong sense that C has no support points.

* Received November 5, 1962.

[1] Preparation of this paper was supported in part by the National Science Foundation, U. S. A. (NSF-GP-378). The author is indebted to R. R. Phelps for some expository suggestions.

To start the construction of C, we choose a convergent series $\sum \alpha_k$ of positive numbers and a continuous concave function δ which takes $[0, \infty[$ onto $[0, 1[$ and is (necessarily) strictly isotone. And we choose also, for some $m \in N$, a closed convex set C_m for which $P \cap L_m \subset C_m \underset{\neq}{\subset} L_m$. (For example, we may choose $\alpha_k = 2^{-k}$, $\delta(\tau) = \tau/(1+\tau)$, and $C_1 = P \cap L_1$.) Having defined $C_m, \cdots, C_i$ $(i \geqq m)$, we define

$$(0) \qquad C_{i+1} = \{x \in L_{i+1} \colon x_{i+1} \geqq 0 \text{ and } \pi_i(x) \in C_i + \alpha_i \delta(x_{i+1}) U_i\},$$

where U_i is the (Euclidean) open unit ball of L_i. (Thus

$$U_i = \{x \in L_i \colon \sum_{j=1}^{i} (x_j)^2 \leqq 1\},$$

and the condition on $\pi_i(x)$ requires the point $(x_1, \cdots, x_i, 0, 0, \cdots)$ to lie in the open $\alpha_i \delta_i(x_{i+1})$-neighborhood of the set C_i.) The set C is then defined as the closure in $R^{\aleph_0}$ of the union $\bigcup_{j=m}^{\infty} C_j$.

We make the following claims:

(1) $C_m \subset C_{m+1} \subset C_{m+2} \subset \cdots$;

(2) *each set* C_i *is convex* $(i \geqq m)$.

In fact, (1) is evident and (2) follows from an easy inductive argument which employs the convexity of the sets C_{i-1} and U_{i-1} and the concavity of the function δ.

Now define $\eta_{ii} = 0$, $\eta_{ij} = \sum_{k=i}^{j-1} \alpha_k$ for $1 \leqq i < j$ and $\eta_i = \lim_{j \to \infty} \eta_{ij}$. We claim:

(3) $\pi_i C_j = C_i + \eta_{ij} U_i \quad (j \geqq i \geqq m)$;

(4) $\pi_i(C_j \cap \{x \colon x_k < \tau\}) \subset C_i + (\eta_{ij} - (1 - \delta(\tau + \eta_{kj}))\alpha_{k-1}) U_i$
$(j \geqq i \geqq m, k \geqq i+1, 0 \leqq \tau < \infty)$.

To prove (3), note first that since $\delta[0, \infty[= [0, 1[$ and $U_l = [0, 1[U_l$, it follows from (0) that $\pi_r C_{r+1} = C_r + \alpha_r U_r$ and $\pi_r C_r = C_r + \eta_{rr} U_r = C_r$ for all $r \geqq m$. Now suppose it is known, for a certain i and j, *that* $\pi_i C_j = C_i + \eta_{ij} U_i$. Then we have

$$\begin{aligned} \pi_i C_{j+1} = \pi_i \pi_j C_{j+1} = \pi_i(C_j + \alpha_j U_j) &= \pi_i C_j + \alpha_j \pi_i U_j \\ = C_i + \eta_{ij} U_i + \alpha_j U_i = C_i &+ (\eta_{ij} + \alpha_j) U_i = C_i + \eta_{i(j+1)} U_i. \end{aligned}$$

Thus (3) follows by mathematical induction.

For (4) it suffices (in view of (1)) to consider the case in which $j > k$. We have

$$\begin{aligned}
\pi_i(C_j \cap \{x : x_k < \tau\}) &= \pi_i\pi_k(C_j \cap \{x : x_k < \tau\}) \\
&\subset \pi_i(\pi_k C_j \cap \{x \in L_k : x_k < \tau\}) \underset{1}{=} \pi_i((C_k + \eta_{kj}U_k) \cap \{x \in L_k : x_k < \tau\}) \\
&\underset{2}{\subset} \pi_i\pi_{k-1}((C_k \cap \{x \in L_k : x_k < \tau + \eta_{kj}\}) + \eta_{kj}U_k) \\
&\underset{3}{\subset} \pi_i(C_{k-1} + \delta(\tau + \eta_{kj})\alpha_{k-1}U_{k-1}) + \eta_{kj}\pi_i U_{k-1} \\
&= \pi_i C_{k-1} + (\delta(\tau + \eta_{kj})\alpha_{k-1} + \eta_{kj})\pi_i U_{k-1} \\
&\underset{4}{=} C_i + \eta_{i(k-1)}U_i + (\eta_{(k-1)j} - (1 - \delta(\tau + \eta_{kj}))\alpha_{k-1})U_i \\
&= C_i + (\eta_{ij} - (1 - \delta(\tau + \eta_{kj}))\alpha_{k-1})U_i,
\end{aligned}$$

where several of the steps are routine, $\underset{1}{=}$ and $\underset{4}{=}$ are based on (3), $\underset{2}{\subset}$ requires an easy verification, and $\underset{3}{\subset}$ depends on (0).

Now let $C = \operatorname{cl} \bigcup_{j=m}^{\infty} C_j$. By (1) and (2), C is convex. From (0) and the choice of C_m it is obvious that $C_i \supset P \cap L_i$ for all $i \geqq m$, and hence $C \supset P$. By (3) it is true that for each $i \geqq m$,

$$\pi_i C \supset \bigcup_{j=i}^{\infty} (C_i + \eta_{ij}U_i) = C_i + \eta_i U_i,$$

where of course $C_i + \eta_i U_i$ is an open proper subset of L_i. Thus to complete the proof, it suffices to show that $\pi_i C \subset C_i + \eta_i U_i$. But this is a consequence of (4), for (4) implies that if $x^1, x^2, \cdots$, is a sequence of points in C such that the sequence $\pi_i(x^1), \pi_i(x^2), \cdots$ converges to a point of the boundary of $C_i + \eta_i U_i$ in L_i, then $\lim_{h\to\infty} x_k^h = \infty$ for all $k \geqq i+1$, where $x^h = (x_1^h, x_2^h, \cdots)$. Thus the set C has no support points even though it is a closed convex proper subset of the complete metrizable locally convex space R.

The question and theorems of Bishop and Phelps [1, 5] referred to supporting hyperplanes corresponding to linear functionals f which are continuous over the entire space. However, the question is also of interest under weaker continuity requirements and in the absence of continuity. Note that *no* nontrivial linear functional (continuous or not) attains a maximum or minimum on the set C constructed above. However, the following questions are apparently open.

(Q1) *If K is a bounded closed convex subset of a complete locally convex space E, must K have a support point? If not by means of a linear*

functional $f \in E^*$, *what about an* f *for which the restriction* f_K *is continuous or even one which is merely algebraically linear with no continuity restriction?*

(Q2) *If* K *is a closed convex subset of a normed linear space, must* K *have a support point by means of a linear functional* f *for which* f_K *is continuous or at least by means of some linear functional? If not in the general case, what happen when* K *is bounded?*

These questions are not answered by the example above (the set C) or by the example in [3] (the set A mentioned earlier), since every bounded subset of $R^{\aleph_0}$ is relatively compact and the set A is supported by hyperplanes corersponding to linear functionals f for which f_A is continuous.

UNIVERSITY OF WASHINGTON.

REFERENCES.

[1] Errett Bishop and R. R. Phelps, "The support functionals of a convex set," *Proceedings of the Symposium on Pure Mathematics*, American Mathematical Society, vol. 7 (Convexity) (1963) to appear.

[2] N. Bourbaki, *Espaces vectoriels topologiques* (Chaps. III-V)„ A. S. I. 1229, Hermann, Paris, 1955.

[3] Victor Klee, "Extremal structure of convex sets II," *Mathematische Zeitschrift*, vol. 69 (1958), pp. 90-104.

[4] G. W. Mackey, "Equivalence of a problem in measure theory to a problem in the theory of vector lattices," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 719-722.

[5] R. R. Phelps, "Support cones and their generalizations," *Proceedings of the Symposium on Pure Mathematics*, American Mathematical Society, vol. 7 (Convexity) (1963) to appear.

THE DIMENSION OF SPACES OF AUTOMORPHIC FORMS.*[1]

By R. P. LANGLANDS.

1. The trace formula of Selberg reduces the problem of calculating the dimension of a space of automorphic forms, at least when there is a compact fundamental domain, to the evaluation of certain integrals. Some of these integrals have been evaluated by Selberg. An apparently different class of definite integrals has occurred in Harish-Chandra's investigations of the representations of semi-simple groups. These integrals have been evaluated. In this paper, after clarifying the relation between the two types of integrals, we go on to complete the evaluation of the integrals appearing in the trace formula. Before the formula for the dimension that results is described let us review Harish-Chandra's construction of bounded symmetric domains and introduce the automorphic forms to be considered.

If $\bar{G}$ is the connected component of the identity in the group of pseudo-conformal mappings of a bounded symmetric domain then $\bar{G}$ has a trivial centre and a maximal compact subgroup of any simple component has non-discrete centre. Conversely if $\bar{G}$ is a connected semi-simple group with these two properties then $\bar{G}$ is the connected component of the identity in the group of pseudo-conformal mappings of a bounded symmetric domain [2(d)]. Let $\mathfrak{g}$ be the Lie algebra of $\bar{G}$ and $\mathfrak{g}_c$ its complexification. Let G_c be the simply-connected complex Lie group with Lie algebra $\mathfrak{g}_c$; replace $\bar{G}$ by the connected subgroup G of G_c with Lie algebra $\mathfrak{g}$. Let K be a maximal compact subgroup of G with Lie algebra $\mathfrak{k}$; then $\mathfrak{k}$ contains a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. Fix once and for all an order on $\mathfrak{h}$. This order is to be so chosen that $\mathfrak{g}_c$ is the direct sum of $\mathfrak{k}_c$, $\mathfrak{p}_+$, and $\mathfrak{p}_-$; $\mathfrak{p}_+$ is spanned by the root vectors belonging to the totally positive roots and $\mathfrak{p}_-$ by the root vectors belonging to the totally negative roots. Moreover $\mathfrak{p}_+$ and $\mathfrak{p}_-$ are abelian and $[\mathfrak{k}_c, \mathfrak{p}_+] \subseteq \mathfrak{p}_+$ and $[\mathfrak{k}_c, \mathfrak{p}_-] \subseteq \mathfrak{p}_-$. Let P_+, P_-, and K_c be the connected subgroups of G_c with Lie algebras $\mathfrak{p}_+$, $\mathfrak{p}_-$, and $\mathfrak{k}_c$ respectively. The exponential mapping of $\mathfrak{p}_+$ into P_+ is bijective; thus P_+ is provided with the structure of a complex vector space. Moreover $G \subseteq P_+K_cP_-$ and $P_+ \cap K_cP_- = \{1\}$. Then $P_+K_cP_-/K_cP_-$ which is identified with $\mathfrak{p}_+$ is a complex vector space and the image of G is a bounded symmetric

* Received November 26, 1962.

[1] Research supported in part by the U. S. Army Research Office (Durham) and in part by the Air Force Office of Scientific Research.

domain B. Finally it should be observed that $G \cap K_c P_- = K$ and that $\mathfrak{p}_+$ is an open subset of the space G_c/K_cP_-. Now identify $\mathfrak{p}_+$ with complex coordinate space and let z be the column of coordinates. If $g \in G_c$, $z \in \mathfrak{p}_+$, and $z' = g(z) \in \mathfrak{p}_+$ (in the space G_c/K_cP_-) let $dz' = \mu(g,z)\,dz$. Before defining the automorphic forms it is necessary to establish a lemma.

LEMMA 1. *Let $\bar{K}_c$ be the restriction of K_c to $\mathfrak{p}_+$ then $\mu(g,z) \in \bar{K}_c$.*

Suppose $X \in \mathfrak{p}_+$ and $f(\cdot)$ is holomorphic in a neighborhood of z' on $\mathfrak{p}_+$ or P_+ since they are identified. Set $h(p_+kp_-) = f(p_+)$, $p_+ \in P_+$, $k \in K_c$, $p_- \in P_-$ then $h(\cdot)$ is a holomorphic function on part of G_c. Let $z = p \in P_+$ and $z' = p' \in P_+$ then, at $t = 0$,

$$\begin{aligned}
\frac{d}{dt}h(g\exp(tX)p) &= \frac{d}{dt}h(\exp(tg(X))p') \\
&= \frac{d}{dt}h(p'\exp(tp'^{-1}g(X))) \\
&= L(X_1)h(p') + L(X_2)h(p') + L(X_3)h(p') \\
&= L(X_1)h(p') - R(X_1)h(p').
\end{aligned}$$

Here $p'^{-1}g(X) = X_1 + X_2 + X_3$, $X_1 \in \mathfrak{p}_+$, $X_2 \in \mathfrak{k}_c$, $X_3 \in \mathfrak{p}_-$. $L(X_i)$ and $R(X_i)$ denote the obvious left or right invariant differential operators. It is necessary to verify that the map $X \to X_1$ is given by an element of K_c. But $gp = p'kp_-$ so $p'^{-1}g = kp_-p^{-1}$ and $kp_-p^{-1}(X) = kp_-(X) \equiv k(X) \pmod{\mathfrak{k}_c + \mathfrak{p}_-}$; thus $X_1 = k(X)$. Finally it should be remarked that $\mu(g,z)$ is a holomorphic function of g and z.

Suppose that σ is an irreducible, holomorphic matrix representation of $\bar{K}_c$ of degree d which is unitary on $\bar{K}$. Then, since $\mu(g_1g_2, z) = \mu(g_1, g_2z)\mu(g_2, z)$, it is easily seen that the action of G on the space $H(\sigma)$ of holomorphic functions on B, whose values are column vectors of length d, defined by $g^{-1}f(z) = \sigma^{-1}(g,z)f(gz)$, with $\sigma(g,z) = \sigma(\mu(g,z))$, is a representation of G. If Γ is a discrete subgroup of G define an (unrestricted) automorphic form of type σ to be a function f in $H(\sigma)$ such that $\gamma f = f$ for all γ in Γ. For subgroups of the symplectic group this definition is essentially the same as that of [7]. As is shown there the dimension of the space, $H(\Gamma,\sigma)$, of automorphic forms of type σ is finite if G/Γ is compact. For a large class of representations σ the calculations of this paper lead to the following formula for the dimension

$$N(\Gamma,\sigma) = \sum_{\{\gamma\}} \nu(G_\gamma/\Gamma_\gamma)\chi(\gamma). \tag{1}$$

The sum is over a set of representatives of those conjugacy classes of Γ that

have a fixed point in B. G_γ is the centralizer of γ in G and Γ_γ is the centralizer of γ in Γ. If the Haar measure ν on G_γ is appropriately normalized then $\chi(\gamma)$ equals

$$(2)\qquad \begin{aligned}&(-1)^{b_\gamma}/v(B_\gamma)\{[G_\gamma\colon G_\gamma{}^0]\prod_{\alpha\in P_\gamma}\rho_\gamma(H_\alpha)\prod_{\substack{\alpha\in P\\ \alpha\notin P_\gamma}}(e^{\frac12\alpha(H)}-e^{-\frac12\alpha(H)})\}^{-1}\\ &\qquad\times\sum_{w_\gamma\backslash w}\epsilon(s)\prod_{\alpha\in P_\gamma}(s\Lambda(H_\alpha)+s\rho(H_\alpha))e^{s\Lambda(H)+s\rho(H)}.\end{aligned}$$

The various symbols will be explained in the course of the proof. This formula agrees with those presented in [3] and [6].

2. In this paragraph and the next the trace formula is reviewed in our special context and a first connection with the work of Harish-Chandra is established. The end result is formula (1) with the numbers $\chi(\gamma)$ expressed as integrals.

Since $\mu(k,z)=\bar{k}$ if $k\in K_0$ the measure

$$dz=|\det(\mu(g,0))|^{-2}\prod_i dx_i dy_i,$$

with $z=g(0)$, is well defined on B and invariant under G. The invariant measure on G is to be so normalized that

$$\int_B f(z)\,dz=\int_G f(g(0))\,dg.$$

Set $G(z)=\sigma^{*-1}(g,0)\sigma^{-1}(g,0)$ with $z=g(0)$. $G(z)$ is well defined and $G(g(z))=\sigma^{*-1}(g,z)G(z)\sigma^{-1}(g,z)$. Introduce the space $H^2(\sigma)$ of functions f in $H(\sigma)$ for which $\int f^*(z)G(z)f(z)\,dz$ is finite. The action of G on $H^2(\sigma)$ is easily seen to be unitary. The functional $f\to f_j(z)$, where $f_j(z)$ is the j-th coordinate of $f(z)$, is bounded on $H^2(\sigma)$; let

$$\int_B g_j{}^*(z_1,z_2)G(z_2)f(z_2)\,dz_2=f_j(z_1)$$

and set $K(z_1,z_2)=(g_1(z_1,z_2),\cdots,g_d(z_1,z_2))^*$. Observe that $K(gz_1,gz_2)=\sigma(g,z_1)K(z_1,z_2)\sigma^*(g,z_2)$. If $L^2(\sigma)$ is the space of measurable functions f on B for which $\int f^*(z)G(z)f(z)\,dz$ is finite then

$$g(z_1)=\int_B K(z_1,z_2)G(z_2)f(z_2)\,dz_2$$

defines the orthogonal projection of $L^2(\sigma)$ onto $H^2(\sigma)$. Consequently

$$K^*(z_2,z_1)=K(z_1,z_2)\quad\text{and}\quad\int_B K(z_3,z_2)G(z_2)K(z_2,z_1)\,dz_2=K(z_3,z_1).$$

Although not necessary it is convenient to verify now that the representation of G in $H^2(\sigma)$ is equivalent to a representation investigated by Harish-Chandra [2(c)]. Let W be the inverse image of B under the map $G_c \to G_c^{-1} \to G_c^{-1}/K_cP_-$ ($W = P_-K_cB^{-1}$ if B is considered a subset of P_+); then if $g \in W$ and $f \in H(\sigma)$ set $f(g) = \sigma^{-1}(g^{-1}, 0)f(g^{-1}(0))$. Then $f(g)$ satisfies: $(\alpha_0) f(pkg) = \sigma(k)f(g)$ if $p \in P_-$ and $k \in K_c$; moreover $f(g)$ is holomorphic on W and if $f(z)$ is in $H^2(\sigma)$ then

$$\| f(\cdot)\|^2 = \int_B f^*(z)G(z)f(z)\,dz = \int_G \| f(g)\|^2\, dg.$$

So the mapping is an isometry on $H^2(\sigma)$. The kernel is replaced by

$$K(g_1, g_2) = \sigma^{-1}(g_1^{-1}, 0)K(g_1^{-1}(0), g_2^{-1}(0))\sigma^{*-1}(g_2^{-1}, 0).$$

Observe that (i) $K(k_1g_1, k_2g_2) = \sigma(k_1)K(g_1, g_2)\sigma^*(k_2)$ if $k_1, k_2 \in K_c$, (ii) $K(pg_1, g_2) = K(g_1, pg_2) = K(g_1, g_2)$ if $p \in P_-$, (iii) $K(g_1g, g_2g) = K(g_1, g_2)$ if $g \in G$, (iv) $K^*(g_2, g_1) = K(g_1, g_2)$, and (v) $\int K(g_1, g_2)K(g_2, g_3)\,dg_2 = K(g_1, g_3)$.

Now we introduce a third space of functions. Suppose $\mathfrak{k}'_c$ is the semi-simple part of $\mathfrak{k}_c$ and $\mathfrak{c}$ is the centre of $\mathfrak{k}_c$ then $\mathfrak{k}_c = \mathfrak{k}'_c + \mathfrak{c}$ and $\mathfrak{h}_c = \mathfrak{h}_c \cap \mathfrak{k}'_c + \mathfrak{c}$. Any linear functional on $\mathfrak{h}_c \cap \mathfrak{k}'_c$ may be extended to $\mathfrak{h}_c$ by setting it equal to zero on $\mathfrak{c}$. Then the given order on the real linear functions on $\mathfrak{h}_c$ induces an order on the real linear functions on $\mathfrak{h}_c \cap \mathfrak{k}'_c$. The representation σ restricted to $\mathfrak{k}'_c$ is irreducible; let ψ_0 be a unit vector belonging to the highest weight with respect to the above order. Then, for all $h \in \mathfrak{h}_c$, $h\psi_0 = \Lambda(h)\psi_0$ where Λ is a linear functional on $\mathfrak{h}_c$. Extending the customary language call Λ the highest weight of σ. If $f(g)$ is a holomorphic function on W satisfying (α_0) above set $h(g) = (f(g), \psi_0)$. Then (α) if $p \in P_-$, $h(pg) = h(g)$; (β) if $n \in N'$, the connected group with Lie algebra $\mathfrak{n}' = \sum \boldsymbol{C}X_{-\alpha}$, the sum being over the positive roots α for which $X_{-\alpha} \in \mathfrak{k}_c$, then $h(ng) = h(g)$; and (γ) if a is in the Cartan subgroup A of G_c with algebra $\mathfrak{h}_c$ then $h(ag) = \xi(a)h(g)$ with $\xi(a) = e^{\Lambda(H)}$ if $a = \exp(H)$. Conversely given a holomorphic function on W satisfying (α), (β), and (γ) there is a holomorphic function $f(g)$ so that $h(g) = (f(g), \psi_0)$. Indeed for fixed g the function $h'(k) = h(kg)$ on K satisfies (i) $h'(ak) = \xi(a)h'(k)$ if $a \in A \cap K$ and (ii) $R(X)h'(k) = 0$ if $X \in \mathfrak{n}'$. Let l be an index for the classes of inequivalent irreducible representations of K and let $(\psi^l_{ij}(k))$ be the matrices of the representations chosen with respect to a basis $(\phi_1, \cdots, \phi_{d_l})$ consisting of eigenvectors of $\mathfrak{h}$; moreover suppose ϕ_1 belongs to the highest weight. Then $h'(k) \sim \sum_l \sum_{i,j} a_{ij}{}^l \psi_{ij}{}^l(k)$. Using (i) and (ii) it is easily seen that, first of all, $a_{ij}{}^l = 0$ unless $i = 1$

and then that $\alpha_{ij}{}^l = 0$ unless $l = l(\sigma)$. So $h'(k) = \sum_j \alpha_{1j}\sigma_{1j}(k)$. Set $f(g) = \sum \alpha_{ij}\phi_j$ then $h(g) = (f(g), \psi_0)$; moreover $f(g)$ is a holomorphic function of g satisfying (α_0) above. Finally the Schur orthogonality relations imply that

$$\begin{aligned}\int_G |h(g)|^2\, dg &= \int_G |(f(g), \psi_0)|^2\, dg \\ &= \int_K dk \int_G |(f(kg), \psi_0)|^2\, dg \\ &= d^{-1} \int_G \| f(g) \|^2\, dg.\end{aligned}$$

This shows that the representation of G on $H^2(\sigma)$ is equivalent to the representation π_Λ studied by Harish-Chandra [2(c)].

Now set $\psi(g_1, g_2) = (K(g_1, g_2)\psi_0, \psi_0)$. This function satisfies (i) $\psi(ng, 1) = \psi(g, 1)$ if $n \in N'$, (ii) $\psi(pg, 1) = \psi(g, 1)$ if $p \in P_-$, (iii) $\psi(ag, 1) = \xi(a)\psi(g, 1)$ if $a \in A$, (iv) $\psi(ga, 1) = \overline{\xi(a)^{-1}}\psi(g, 1)$ if $a \in A \cap K$, and (v) $\psi(g, 1)$ is a holomorphic function on W. But Harish-Chandra ([2(c)], p. 22) has shown that there is essentially only one function with these properties so $\psi(g, 1) = \delta\psi_\Lambda(g)$. $\delta = \psi(1, 1)$ and $\psi_\Lambda(g) = (\zeta(g)\phi_0, \phi_0)e^{\lambda(\Gamma(g))}$. Here ζ is a representation of G_c with highest weight Λ_0 and $\lambda = \Lambda - \Lambda_0$; moreover Λ_0 is so chosen that λ vanishes on $\mathfrak{h}_c \cap \mathfrak{k}'_c$. ϕ_0 is a unit vector belonging to the weight Λ_0. The function $\mu^{-1}(g^{-1}, 0)$ is a holomorphic function on W with values in K_c. It may be lifted to a function on $\tilde{W}$, the universal covering space of W, with values in $\tilde{K}_c$, the universal covering group of K_c. $\tilde{K}_c$ is the product of a simply connected, complex abelian group C with Lie algebra $\mathfrak{c}$ and a semi-simple group. Mapping $\tilde{W}$ into $\tilde{K}_c$, projecting on C, and then taking the logarithm one obtains $\Gamma(g)$ which lies in $\mathfrak{c}$. Thus $\Gamma(g)$ is a single-valued function on $\tilde{W}$ but a multiple-valued function on W.

Certainly $\delta \neq 0$ if $H^2(\sigma) \neq \{0\}$. In particular, if X_β is a root vector belonging to the positive root β, if $X_{-\beta}$ belongs to $-\beta$ and $H_\beta = [X_\beta, -X_{-\beta}]$ then ([2(d)], p. 612) $H^2(\sigma) \neq \{0\}$ if $2\beta^{-1}(H_\beta)(\Lambda(H_\beta) + \rho(H_\beta)) < 0$ for every totally positive root β. ρ is one-half the sum of the positive roots.

3. It will now be supposed that $2\beta^{-1}(H_\beta)(\Lambda(H_\beta) + \rho(H_\beta) + 2\rho_+(H_\beta)) < 1$ for every totally positive root β; ρ_+ is one-half the sum of the totally positive roots. Then ([2(d)], p. 610) $\psi_\Lambda(g)$ is integrable and, since σ is irreducible, $K(g, 1)$ is integrable. Let $H^\infty(\sigma)$ be the space of functions in $H(\sigma)$ such that $f^*(z)G(z)f(z)$ is bounded. Then, if $f(z)$ is in $H^\infty(\sigma)$,

$$\int_B K(z_1, z_2)G(z_2)f(z_2)\, dz_2$$

converges. To verify that it equals $f(z_1)$ it is sufficient to show that $H^2(\sigma)$ contains all polynomials for then the argument of Godement in [7] applies. To do this it is sufficient to show that $G(z)$ is integrable over B. This is the same as showing that $\|\sigma^{-1}(g^{-1},0)\|^2$ is integrable over G. Let A_+ be the connected group with Lie algebra $\mathfrak{a}_{\mathfrak{p}_0}$ ([2(d)], p. 583) then every element of G may be written as k_1ak_2 with $k_1, k_2 \in K$ and $a \in A_+$. Moreover

$$\|\sigma^{-1}(g^{-1},0)\|^2 = \|\sigma^{-1}(a^{-1},0)\| = \|\sigma(h(a))\|^2$$

([2(d)], p. 599). But $\|\sigma(h(a))\|^2 = \operatorname{tr}(\sigma(h^2(a)))$ which is known to be integrable. Let $L^2(\Gamma,\sigma)$ be the space of measurable functions on B, whose values are column vectors of length d, such that $\sigma^{-1}(\gamma,z)f(\gamma z) = f(z)$ for all z in B and all γ in Γ and $\int_F f^*(z)G(z)f(z)\,dz$ is finite, where F is a fundamental domain for Γ in B. Then

$$f(\cdot) \to \int_B K(\cdot,z)G(z)f(z)\,dz$$

defines the orthogonal projection of $L^2(\Gamma,\sigma)$ onto $H(\Gamma,\sigma)$. Now

$$\int_B K(z_1,z_2)G(z_2)f(z_2)\,dz_2 = \int_F \sum_\Gamma K(z_1,\gamma z_2)\sigma^{*-1}(\gamma,z_2)G(z_2)f(z_2)\,dz_2$$

provided $\sum_\Gamma K(z_1,\gamma z_2)\sigma^{*-1}(\gamma,z_2)$ is uniformly absolutely convergent for z_1 and z_2 in F. To verify this it is sufficient to show that $\sum_\Gamma K(g_1,g_2\gamma)$ converges uniformly absolutely in some neighborhood of each point (g'_1,g'_2) in $G\times G$. Since σ is irreducible it is enough to consider the series $\sum_\Gamma |\psi_\Lambda(g_1\gamma g_2^{-1})|$. Writing $g = k_1ak_2$ we have ([2(d)], pp. 598-600)

$$\begin{aligned}|\psi_\Lambda(g)| &\leqq |(\zeta(k_1^{-1})\phi_0, \zeta(a)\zeta(k_2)\phi_0)e^{\lambda(\Gamma(a))}| \\ &\leqq e^{\lambda(\Gamma(a))}\chi_{\Lambda_0}(h(a)).\end{aligned}$$

Let ϕ be the mapping of G onto the symmetric space G/K and let r be the metric on G/K. Every element of G may be written as a product $g = k_1ak_2$ with $k_1, k_2 \in K$ and with $a \in A_+$ such that $\log a \in \mathfrak{a}_{\mathfrak{p}_0}{}^+ = \{X \in \mathfrak{a}_{\mathfrak{p}_0} \mid \alpha(X) \geqq 0$ for all positive roots $\alpha\}$. To be more precise one introduces an order on the linear functions on $\mathfrak{a}_{\mathfrak{p}_0}$, extends $\mathfrak{a}_{\mathfrak{p}_0}$ to a Cartan subalgebra of $\mathfrak{g}$, extends the ordering, and takes the positive roots of this subalgebra with respect to the resulting order. Although it is not *a priori* uniquely determined by g we set $a(g) = a$. We want to show that if $\epsilon > 0$ is given it is possible to choose ϵ_1 so that if h_1 and h_2 are in G and $r(\phi(h_1),\phi(h_2)) < \epsilon_1$ then $|\nu(\log(a(h_1)) - \log(a(h_2)))|$

$\leqq \epsilon \| \nu \|$ for any linear functional ν on $\mathfrak{a}_{\mathfrak{p}_0}$. It is enough to establish this for a basis of the space of linear functionals which may be supposed to consist of the highest weights of certain representations of G restricted to $\mathfrak{a}_{\mathfrak{p}_0}$. Let π be such a representation which may be supposed to satisfy $\pi(\theta(g)) = \pi^{*-1}(g)$ if θ is the Cartan involution of G leaving K fixed. Then $g \rightarrow \pi(g)\pi^*(g)$ defines an imbedding of G/K in a manner which we shall pretend is isometric in the space of positive definite Hermitian matrices with the Riemannian metric $d^2Y = \operatorname{tr}(Y^{-1}dYY^{-1}dY)$. So it is enough to show that if P_1 and P_2 are positive definite matrices with maximum eigenvalues λ_1, λ_2 then $|\log \lambda_1/\lambda_2| < \epsilon$ if $r(P_1, P_2) < \epsilon$ (cf. [2(i)], p. 280). If $P_1 = AA^*$ and $P_2 = A(\delta_{ij}e^{\alpha_i})A^*$, $\alpha_1 \geqq \alpha_2 \geqq \cdots \geqq \alpha_n$, then $r(P_1, P_2) = (\sum_{i=1}^{n} \alpha_i^2)^{\frac{1}{2}}$. Let $\| x \| = 1$ and $\| A^*x \|^2 = \lambda_1$; then, if $y = A^*x$, $\lambda_2 \geqq \sum_{i=1}^{n} e^{\alpha_i} y_i^2 \geqq e^{\alpha_n}\lambda_1$. Similarly $\lambda_1 \geqq e^{-\alpha_1}\lambda_2$; so $-\alpha_1 \leqq \log \lambda_1/\lambda_2 \leqq -\alpha_n$.

As a consequence there are positive numbers c and $\delta < 1$ so that if $U(g') = \{g \mid r(\phi(g), \phi(g')) < \delta\}$ then

$$e^{\lambda(\Gamma(a(g')))}\chi_{\Lambda_0}(h(a(g'))) \leqq c \int_{U(g')} e^{\lambda(\Gamma(a(g)))}\chi_{\Lambda_0}(h(a(g)))\,dg.$$

Let U_1 and U_2 be compact neighborhoods of g'_1 and g'_2 respectively. There is an integer N so that for $g_1 \in U_1$ and $g_2 \in U_2$ any point in G belongs to at most N of the sets $U(g_1\gamma g_2^{-1})$, $\gamma \in \Gamma$. Finally given $\epsilon > 0$ there is a positive number M so that if $V_M = \{g \mid r(\phi(g), \phi(1)) \geqq M\}$ then

$$\int_{V_M} e^{\lambda(\Gamma(a(g)))}\chi_{\Lambda_0}(h(a(g)))\,dg < \epsilon.$$

For all but a finite set Γ_1 of elements of Γ, $U_1\gamma U_2^{-1} \subseteq G - V_{M+1}$. Thus

$$\sum_{\gamma \notin \Gamma_1} \exp(\lambda(\Gamma(a(g_1\gamma g_2^{-1}))))\chi_{\Lambda_0}(h(a(g_1\gamma g_2^{-1}))) \leqq cN\epsilon$$

which was to be shown.

The kernel $\sum_{\Gamma} K(z_1, \gamma z_2)\sigma^{*-1}(\gamma, z_2)$ is continuous; thus if $\{\omega^i\}$, $i = 1, \cdots, N(\Gamma, \sigma)$, is an orthonormal basis for $H(\Gamma, \sigma)$

$$\sum K(z_1, \gamma z_2)\sigma^{*-1}(\gamma, z_2) = \sum_{i=1}^{N(\Gamma,\sigma)} \omega^i(z_1)\omega^{i*}(z_2).$$

But

$$N(\Gamma, \sigma) = \sum_{i=1}^{N(\Gamma,\sigma)} \int_F \sum_{j,l} \bar{\omega}_l^i(z) \; G_{lj}(z)\omega_j^i(z)\,dz;$$

consequently

$$N(\Gamma, \sigma) = \int_F \operatorname{tr}\{\sum_{\gamma} K(z, \gamma z)\sigma^{*-1}(\gamma, z)G(z)\}\,dz.$$

Following Selberg [5] this may be written

$$\sum_{\{\gamma\}}\sum_{\Gamma_\gamma\backslash\Gamma}\int_F \operatorname{tr}\{K(z,\beta^{-1}\gamma\beta z)\sigma^{*-1}(\beta^{-1}\gamma\beta,z)G(z)\}dz.$$

The outer sum is over a set of representatives of the conjugacy classes of Γ; the inner sum is over a set of coset representatives of the centralizer Γ_γ of γ in Γ. Rewrite the last sum as

$$\sum_{\{\gamma\}}\sum_{\Gamma_\gamma\backslash\Gamma}\int_F \operatorname{tr}\{\sigma^{-1}(\beta,z)K(\beta z,\gamma\beta z)\sigma^{*-1}(\gamma\beta,z)G(z)\}dz$$
$$=\sum_{\{\gamma\}}\int_{F_\gamma}\operatorname{tr}\{K(z,\gamma z)\sigma^{*-1}(\gamma,z)G(z)\}dz.$$

F_γ is a fundamental domain for Γ_γ in B. Replace these integrals by integrals over fundamental domains F'_γ for Γ_γ acting on G to the right to obtain

$$\sum_{\{\gamma\}}\int_{F'_\gamma}\operatorname{tr}\{K(g^{-1}(0),\gamma g^{-1}(0))\sigma^{*-1}(\gamma,g^{-1}(0))G(g^{-1}(0))\}dg$$
$$=\sum_{\{\gamma\}}\int_{F'_\gamma}\operatorname{tr}\{K(g\gamma g^{-1},1)\}dg.$$

According to [5] this equals

$$\sum_{\{\gamma\}}\nu(G_\gamma/\Gamma_\gamma)\int_{S_\gamma}\operatorname{tr}\{K(g\gamma g^{-1},1)\}ds_\gamma;$$

G_γ is the centralizer of γ in G; $S_\gamma=G/G_\gamma$; and the measures are so normalized that $dg=ds_\gamma dg_\gamma$. $\nu(G_\gamma/\Gamma_\gamma)$ is the volume of a fundamental domain for Γ_γ acting to the left (or right) in G_γ. It is of some importance to observe that every integral appearing is absolutely convergent. If γ is in Γ set

$$\chi(\gamma)=\int_{S_\gamma}\operatorname{tr}\{K(g\gamma g^{-1},1)\}ds_\gamma;$$

apart from the arbitrariness of the invariant measure on S_γ, $\chi(\gamma)$ depends only on the conjugacy class of γ in G. By the Schur orthogonality relations

$$\chi(\gamma)=d\int_{S_\gamma}\int_K(K(kg\gamma g^{-1}k^{-1},1)\psi_0,\psi_0)\,dk\,ds_\gamma$$
$$=d\delta\int_{S_\gamma}\int_K\psi_\Lambda(kg\gamma g^{-1}k^{-1})\,dk\,ds_\gamma.$$

4. The first step in the evaluation of these integrals is to calculate $d\delta$. Now

$$\|(K(g,1)\psi_0,\psi_0)\|^2 = \int_G |(K(g,1)\psi_0,\psi_0)|^2\,dg$$
$$= \int_G \int_K |(\sigma(k)K(g,1)\psi_0,\psi_0)|^2\,dk\,dg$$
$$= d^{-1}\int_G \|K(g,1)\psi_0\|^2\,dg$$
$$= d^{-1}\int_G |K^*(g,1)K(g,1)\psi_0,\psi_0)\,dg$$
$$= d^{-1}(K(1,1)\psi_0,\psi_0) = d^{-1}\delta$$

which shows that $d\delta = \|\psi_\Lambda(g)\|^{-2}$ since $\|(K(g,1)\psi_0,\psi_0)\|^2 = \delta^2\,\|\psi_\Lambda(g)\|^2$. But $\|\psi_\Lambda(g)\|^{-2}$ has been calculated by Harish-Chandra ([2(d)], p. 608); it equals

$$c(G)\prod_{\beta\in P}|\Lambda(H_\beta)+\rho(H_\beta)/\rho(H_\beta)|$$

where P is the set of positive roots and $c(G)$ is a constant independent of Λ. To calculate $c(G)$ take $\sigma(\bar{k}) = (\det\bar{k})^{-1}$, $\bar{k}\in\bar{K}_0$, so that $\Lambda = -2\rho_+$; this is permissible since, as will be seen in a moment, $2\beta^{-1}(H_\beta)(-2\rho_+(H_\beta)+\rho(H_\beta)) < 0$ for every totally positive root β. Then

$$\|\psi_\Lambda(g)\|^{-2} = c(G)\prod_{\beta\in P}|-2\rho_+(H_\beta)+\rho(H_\beta)/\rho(H_\beta)| = c(G)$$

since $\prod_{\beta\in P}|-2\rho_+(H_\beta)+\rho(H_\beta)| = \prod_{\beta\in P}|\rho(H_\beta)|$. To see this observe that ([2(b)], p. 749) one could choose as a set of positive roots the positive roots with root vectors in $\mathfrak{k}_0$ and the negatives of the totally positive roots. Let ρ' be one-half the sum of the positive roots in this new order. Then $\rho' = \rho - 2\rho_+$ and $2\beta^{-1}(H_\beta)\rho'(H_\beta) = -2(-\beta(H_{-\beta}))^{-1}\rho'(H_{-\beta}) < 0$ since $-\beta$ is positive in this new order if β is totally positive. There is an element s in the normalizer of $\mathfrak{h}_0$ in G_0 which takes the positive roots in the original order into the positive roots in the new order; in particular $\rho'(H) = \rho(s^{-1}(H))$. Now

$$[H,s(X_\alpha)] = s([s^{-1}(H),X_\alpha]) = \alpha(s^{-1}(H))s(X_\alpha) = s(\alpha)(H)s(X_\alpha)$$

and

$$H'_{s(\alpha)} = [X'_{s(\alpha)},X'_{-s(\alpha)}] = [s(X_\alpha),s(X_{-\alpha})] = s(H_\alpha).$$

Now $[H'_{s(\alpha)},X'_{s(\alpha)}] = s[H_\alpha,X_\alpha] = \alpha(H_\alpha)X'_{s(\alpha)}$ so if all X_α are so normalized that $\alpha(H_\alpha) = 2$ then $H'_{s(\alpha)} = H_{s(\alpha)}$. Consequently

$$|\prod_{\beta\in P}\rho'(H_\beta)| = |\prod_{\beta\in P}\rho'(H'_{s(\beta)})| = |\prod_{\beta\in P}\rho'(s(H_\beta))| = |\prod_{\beta\in P}\rho(H_\beta)|.$$

On the other hand $\psi(g) = \det(\mu(g^{-1},0))$ satisfies (i) $\psi(pg) = \psi(g)$ if

$p \in P_-$, (ii) $\psi(gk) = \psi(kg) = \det^{-1}(\bar{k})\psi(g)$ if $k \in K_0$, (iii) $\psi(g)$ is holomorphic on W, and (iv) $\psi(1) = 1$. This is enough to ensure that $\psi(g) = \psi_\Lambda(g)$. Thus

$$\begin{aligned}\| \psi_\Lambda(g) \|^2 &= \int_G |\psi(g)|^2 \, dg \\ &= \int_B |\det(\mu(g,0))|^2 \, |\det(\mu(g,0))|^{-2} \prod_i dx_i dy_i \\ &= v(B).\end{aligned}$$

$v(B)$ is the Euclidean volume of B. In conclusion

$$\chi(1) = (-1)^b / v(B) \prod_{\beta \in P} (\Lambda(H_\beta) + \rho(H_\beta)/\rho(H_\beta)),$$

with b equal to the complex dimension of B.

It will be useful at this point to establish some notation. The universal covering groups of G and K_0 have been denoted by $\tilde{G}$ and $\tilde{K}_0$. If G_1 is a subgroup of G then $\tilde{G}_1$ is the group of all elements in $\tilde{G}$ lying over G_1. Elements of $\tilde{G}$ will be denoted by g and γ and their projection in G by g and γ; similarly $k \in \tilde{K}_0$ projects on k. If N is a simply-connected subgroup of G then N is isomorphic to the connected component of the identity in $\tilde{N}$ so the same symbol will be used for corresponding elements in the two groups. Finally $g\gamma g^{-1}$ will be written $g\gamma g^{-1}$.

Every element of Γ is semi-simple [1]; this implies in particular that $G_\gamma{}^0$, the connected component of the identity in G_γ, is of finite index in G_γ. The measure on G_γ will be so normalized that, on $G_\gamma{}^0$, $dg_\gamma = dg_\gamma{}^0$. Then if the measure on $S_\gamma{}^0 = G/G_\gamma{}^0$ is normalized in the usual manner

$$\chi(\gamma) = d_\Lambda [G_\gamma : G_\gamma{}^0]^{-1} \int_{S_\gamma{}^0} \int_K \psi_\Lambda(kg\gamma g^{-1}k^{-1}) \, dk \, ds_\gamma{}^0$$

with $d_\Lambda = d\delta$. But ψ_Λ may be lifted to a function on $\tilde{G}$ and $\chi(\gamma)$ may be written as

$$d_\Lambda [G_\gamma : G_\gamma{}^0]^{-1} \int_{S_\gamma{}^0} \int_K \psi_\Lambda(kg\gamma g^{-1}k^{-1}) \, dk \, ds_\gamma{}^0. \tag{3}$$

Recall that $\psi_\Lambda(g) = (\zeta(g)\phi_0, \phi_0) e^{\lambda(\Gamma(g))}$. Revising the notation slightly denote the linear functions λ and Λ associated to the representation σ by λ' and Λ' and let λ be an arbitrary linear function on $\mathfrak{h}_c$ vanishing on $\mathfrak{h}_c \cap \mathfrak{k}'_c$ and, accordingly, let $\Lambda = \Lambda_0 + \lambda$. d_Λ and $\psi_\Lambda(g)$, but not $\psi_\Lambda(g)$, are still defined. Let us now see for which functions, λ, the integral converges.

The function $\mu(g, z)$ on $G \times B$ may be lifted to a function $\mu(g, z)$ on $\tilde{G} \times B$ with values in $\tilde{K}_c$ which satisfies $\mu(g_1, g_2 z)\mu(g_2, z) = \mu(g_1 g_2, z)$.

Perhaps the simplest way to see this is to observe that if $z = p$, $p \in P_+$, then p^{-1} is in $\tilde{W} = P_-\tilde{K}_0 B^{-1}$ ([2(c)], p. 5) and so is $p^{-1}g^{-1} = p_- k^{-1} p_+^{-1}$, $p_+ \in B$; then $\mu(g,z) = k$. In particular $\mu(k_1 g k_2, 0) = k_1 \mu(g,0) k_2$ so that

$$\Gamma(k_1 g k_2) = \Gamma(k_1) + \Gamma(g) + \Gamma(k_2).$$

Now write $g = k_1 a k_2$ with $a = \exp(\sum_{i=1}^{s} t_i (X_{\gamma_i} + X_{-\gamma_i}))$ ([2(d)], p. 599). It is possible to choose a basis $\{c_1, \cdots, c_k\}$ for $\mathfrak{c}$ so that the coordinates of $\Gamma(k_1)$ and $\Gamma(k_2)$ are purely imaginary and those of $\Gamma(a)$ are of the form $\sum_i \log(\cosh t_i) a_{ij}$ with $a_{ij} \geqq 0$. If the basis is chosen from $i(\mathfrak{c} \cap \mathfrak{k})$ the first condition is satisfied. The second will be satisfied if we choose a basis so that the projection of H_{γ_i}, $i = 1, \cdots, s$, on the centre has positive coordinates ([2(d)], p. 600). It will be enough to show that this can be done when the group is simple and $\mathfrak{c}$ has dimension 1. But $2\rho_+(H_{\gamma_i}) > 0$, $i = 1, \cdots, s$, and $2\rho_+(H)$ is determined solely by the projection of H on $\mathfrak{c}$ since it is the trace of the representation of $\mathfrak{k}_0$ on $\mathfrak{p}_+$. Since $H_{\gamma_i} \in i\mathfrak{k}$ the assertion is proved. It will be shown below that for fixed γ the imaginary parts of the coordinates of $\Gamma(g\gamma g^{-1})$ remain bounded as g varies over G; consequently the integral over S_γ^0 converges absolutely if $\mathrm{Re}\{\lambda(c_i) - \lambda'(c_i)\} \leqq 0$ and represents a function of λ which is continuous on this set and holomorphic in its interior. Thus it will be sufficient to evaluate the integral (3) when $\lambda(c_i)$ is real and very much less than zero.

We now establish the improved assertion. The notation of [2(d)] will be used. In showing that the imaginary parts of the coordinates of $\Gamma(g\gamma g^{-1})$ are bounded we may suppose that $g = k_1 \exp(X) k_2$ with $X \in \mathfrak{a}_{\mathfrak{p}_0}$ and k_1 and k_2 in some fixed compact subset of $\tilde{K}$. Suppose $\pi(k)$ is the result of projecting k on the centre of $\tilde{K}_0$ and then taking the logarithm. Then

$$\begin{aligned}\Gamma(g\gamma g^{-1}) &= \pi(\mu^{-1}(\exp(X) k_2 \gamma^{-1} k_2^{-1} \exp(-X), 0) \\ &= -\pi(\mu(\exp(-X), 0)) - \pi(\mu(\gamma, k_2^{-1}\exp(-X)(0))) \\ &\qquad -\pi(\mu(\exp(X), k_2\gamma^{-1}k_2^{-1}\exp(-X)(0))).\end{aligned}$$

The first term gives no contribution to the imaginary part. $\mu(\gamma, z)$ is defined for z in an open subset of $\mathfrak{p}_+$ containing the closure of B so it is possible to define $\mu(\gamma, z)$ on the same set. Since it is continuous it takes the closure of B into a compact set. Thus only the third term causes trouble. So we consider $\mu(\exp(X), z)$ letting z vary over B.

The calculations will be simplified if we first prove a lemma. Every element X of $\mathfrak{p}_+$ determines a linear transformation from $\mathfrak{p}_-$ to $\mathfrak{k}_0$, namely,

$T(X)Y = [X, Y]$ if $Y \in \mathfrak{p}_-$. Introduce on $\mathfrak{p}_-$ and $\mathfrak{k}_c$ the Hermitian inner product $-B(Y_1, \tilde{\theta}(Y_2))$ then

LEMMA 2. *B is the set of vectors X in $\mathfrak{p}_+$ for which $2I - T^*(X)T(X)$ is positive definite.*

If k is in K then

$$T(k(X)) = \mathrm{Ad}(k)T(X)\mathrm{Ad}(k^{-1}) \quad \text{and}$$

$$T^*(k(X))T(k(X)) = \mathrm{Ad}(k)T^*(X)T(X)\mathrm{Ad}(k^{-1});$$

moreover X is in B if and only if $k(X)$ is in B. So in proving the lemma we may replace X by any element equivalent to it under the adjoint action of K. Suppose X is in B then X may be supposed equal to $\sum_{i=1}^{s} a_i X_{\gamma_i}$ with $-1 < a_i < 1$. Any element of $\mathfrak{p}_-$ may be written as

$$Y = \sum_{i=1}^{s} b_i X_{-\gamma_i} + \sum_{i} \sum_{\alpha \in P_i} b_\alpha X_{-\alpha} + \sum_{i<j} \sum_{\alpha \in P_{ij}} b_\alpha X_{-\alpha}.$$

Then

$$[X, Y] = \sum_{i=1}^{s} a_i b_i [X_{\gamma_i}, X_{-\gamma_i}] + \sum_{i} \sum_{\alpha \in P_i} a_i b_\alpha [X_{\gamma_i}, X_{-\alpha}] + \sum_{i,j} \sum_{\alpha \in P_{ij}} a_i b_\alpha [X_{\gamma_i}, X_{-\alpha}].$$

It is easily seen that $B(X_\alpha, \tilde{\theta}(X_\beta)) = 0$ unless $\alpha = \beta$ and that $[X_{\gamma_i}, Y]$ is orthogonal to $[X_{\gamma_j}, Y]$ if $i \neq j$. Moreover $\tilde{\theta}([X_{\gamma_i}, X_{-\alpha}]) = [X_{-\gamma_i}, X_\alpha]$ and

$$\begin{aligned} -B([X_{\gamma_i}, X_{-\alpha}], [X_{-\gamma_i}, X_\beta]) &= -B([[X_{\gamma_i}, X_{-\alpha}], X_{-\gamma_i}], X_\beta) \\ &= -B([H_{\gamma_i}, X_{-\alpha}], X_\beta) \\ &= \alpha(H_{\gamma_i}) B(X_{-\alpha}, X_\beta). \end{aligned}$$

Since $\alpha(H_{\gamma_i}) = 0$, 1, or 2 it follows that $\| T(X)Y \|^2 < 2 \| Y \|^2$. Conversely suppose $X \in \mathfrak{p}_+$ and $\| T(X)Y \|^2 < 2 \| Y \|^2$ for every Y in $\mathfrak{p}_-$. If $X = \sum_{i=1}^{s} a_i X_{\gamma_i}$ with a_i real, as may be assumed, then $\| [X, X_{-\gamma_i}] \|^2 = 2a_i^2 \| X_{-\gamma_i} \|^2$ so that $| a_i | < 1$. It follows that X is in B.

Similar calculations now show that if

$$X_0 = \sum_{i=1}^{s} a_i X_{\gamma_i} + \sum_{i} \sum_{\alpha \in P_i} b_\alpha X_\alpha + \sum_{i<j} \sum_{\alpha \in P_{ij}} b_\alpha X_\alpha$$

is in B then $| a_i | < 1$, $i = 1, \cdots, s$.

The original assertion will be proved if we show that the imaginary coordinates of $\pi(\mu(\exp(t(X_{\gamma_i} + X_{-\gamma_i})), z))$ remain bounded as z varies over

B. Let $z = \exp(X_0)$ with X_0 as above and set $g(t) = \exp(t(X_{\gamma_i} + X_{-\gamma_i}))$. Write

$$X_0 = \sum_{\alpha \in S_i} a_\alpha X_\alpha + a_i X_{\gamma_i} + \sum_{\substack{\alpha \notin S_i \\ \alpha \neq \gamma_i}} a_\alpha X_\alpha = X_1 + X_2 + X_3$$

where S_i is the set of roots which vanish on H_{γ_i}. Then $g(t)\exp(X_0) = \exp(X_1)g(t)\exp(X_2)\exp(X_3)$. $g(t)$ and $\exp(X_2)$ belong to the complex group whose Lie algebra is spanned by H_i, X_{γ_i}, and $X_{-\gamma_i}$. A simple calculation in $SL(2;C)$ shows that $g(t) = \exp(a(t)X_{\gamma_i})\exp(b(t)H_{\gamma_i})\exp(c(t)X_{-\gamma_i})$ with

$$a(t) = (a_i \cosh t + \sinh t)(a_i \sinh t + \cosh t)^{-1}$$
$$b(t) = -\log(a_i \sinh t + \cosh t)$$
$$c(t) = \sinh t (a_i \sinh t + \cosh t)^{-1}.$$

Finally

$$\exp(c(t)X_{-\gamma_i})\exp(X_3) = \exp(X_3)\exp(X_4)$$

with $X_4 = \mathrm{Ad}(\exp(-X_3))(c((t)X_{-\gamma_i})$ so $X_4 = c(t)X_{-\gamma_i} + \sum c_\alpha X_\alpha$. The sum is over the positive compact roots. This implies that $\exp(X_4)$ is the product of an element in K'_c, the semi-simple component of $\bar{K}_c$, and an element in P_-. So

$$\pi(\mu(g(t),z)) = \pi(\exp(b(t)H_{\gamma_i})) = -\log(a_i \sinh t + \cosh t)\pi(\exp(H_{\gamma_i})).$$

The coordinates of $\pi(\exp(H_{\gamma_i}))$ are real and, since $|a_i| < 1$,

$$\mathrm{Re}(a_i \sinh t + \cosh t) > 0 \text{ so } -\tfrac{\pi}{2} < \mathrm{Im}(\log(a_i \sinh t + \cosh t)) < \tfrac{\pi}{2}.$$

It will be seen that for $\lambda(c_i) \ll 0$ and γ semi-simple the double integral in (3) is absolutely convergent; consequently in our analysis the integral over K may be omitted and γ need not belong to Γ.

5. It will be convenient in the evaluation of the integrals (3) to omit at first any detailed estimates. These will be discussed in the next paragraph. Suppose that γ is a regular element in G and let γ belong to the centralizer B of the Cartan subalgebra $\mathfrak{j}$ of $\mathfrak{g}$. According to [2(e)] it may be supposed that $\theta(\mathfrak{j}) = \mathfrak{j}$. Hence $\mathfrak{j} = \mathfrak{j}_1 + \mathfrak{j}_2$ with $\mathfrak{j}_1 = \mathfrak{j} \cap \mathfrak{k}$ and $\mathfrak{j}_2 = \mathfrak{j} \cap \mathfrak{p}$. The case that $\mathfrak{j}_2 = \{0\}$ will be treated first. Then $B \subseteq K$ and it may be supposed that $\mathfrak{j} = \mathfrak{h}$. $G_\gamma \subseteq K$ and the measure on G_γ is so normalized that the total measure of $G_\gamma{}^0$ is 1. The integration over $S_\gamma{}^0$ in (3) may then be replaced by an integration over G; as will be seen below the integral is then a continuous function of γ as γ varies over the regular elements in $\bar{K}$. Harish-Chandra has shown that if T_Λ is the character of the representation π_Λ then

$$T_\Lambda(f) = d_\Lambda \int_G dg\{\int_{\bar{G}} f(g_1)\psi_\Lambda(gg_1g^{-1})\,dg_1\}$$

when f is an infinitely differentiable function with compact support. If the support of f is contained in the set of regular elements in GKG^{-1} the order of integration may be reversed. Another formula for $T_\Lambda(f)$ is implicit in the papers [2(c)] and [2(e)]. However before introducing this it must be observed that $\tilde{B}$ is connected and thus every element of $\tilde{B}$ can be written as the exponential of an element in $\mathfrak{h}$. In particular, let $\tilde{\gamma} = \exp(H)$. Then $T_\Lambda(f)$ is obtained by integrating f against a continuous function whose value at γ is

$$\{\prod_{\alpha\in P}(e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})\}^{-1}\sum_{s\in w}\epsilon(s)\,e^{s\Lambda(H)+s\rho(H)}.$$

P is the set of positive roots; w is the Weyl group of K_c; and $\epsilon(s) = \pm 1$ according as s is the product of an even or odd number of reflections. It should be observed that the second hypothesis of Section 10 of [2(e)] does not hold here so it is necessary to prove Lemma 43 using Fourier integrals rather than series. The value of $\chi(\gamma)$ obtained agrees with (2) since B_γ is reduced to a point with $v(B_\gamma) = 1$, P_γ is empty, and w_γ is reduced to $\{1\}$.

Retaining the assumption that γ is regular it will now be supposed that $\mathfrak{j}_2 \neq \{0\}$. Define the subgroups M and N as on page 212 of [2(g)] with $\mathfrak{j}$ replacing $\mathfrak{h}_0$ then ([2(g)], p. 216)

$$\int_{G/G_\gamma^0}\psi_\Lambda(g\gamma g^{-1})\,ds_\gamma^0 = \int_{K\times M\times N}\psi_\Lambda(kmn\gamma n^{-1}m^{-1}k^{-1})\,dk\,dm\,dn,$$

if the Haar measures on M and N are suitably normalized. It should be observed that, contrary to the assertion in [2(e)], the centralizer $\tilde{B}$ in $\tilde{G}$ of a Cartan subalgebra is not always commutative. Thus, if γ belong to $\tilde{B}$ one must consider $\int_{G/B^0} f(g\gamma g^{-1})\,d\bar{g}$ and not $\int_{G/B} f(g\gamma g^{-1})\,d\bar{g}$; B is the projection of $\tilde{B}$ on G and B^0 is the connected component of the identity in B. The theorems of [2(h)] used later must be interpreted with this observation in mind. It is not difficult (cf. [2(a)], p. 509) to see that the above integral equals

$$\xi(X_\iota)^{-1}(\gamma)\int_{K\times M\times N}\psi_\Lambda(knm\gamma m^{-1}k^{-1})\,dk\,dm\,dn$$

with $\xi(X_\iota)(\gamma)$ equal to the determinant of the restriction of $I - ad(\gamma)$ to $\mathfrak{n}$, the Lie algebra of N. It can be assumed that $\mathfrak{j}_2$ is contained in $\mathfrak{a}_{\mathfrak{p}_0}$. Then, in the notation of [2(d)], for some l either $\nu^{-1}(X_{\gamma_l})$ or $\nu^{-1}(X_{-\gamma_l})$ is in $\mathfrak{n}_0$. Since the order on $\mathfrak{j}_2$ is arbitrary suppose that $\nu^{-1}(X_{\gamma_l})$ is in $\mathfrak{n}_0$.

$$\nu^{-1}(X_{\gamma_l}) = \tfrac{1}{2}(X_{\gamma_l} - X_{-\gamma_l} - H_{\gamma_l}) \text{ and } 2i\nu^{-1}(X_{\gamma_l}) = X$$

is in $\mathfrak{g}$ and thus in $\mathfrak{n}$. Let $N_1 = \{\exp(tX) \mid -\infty < t < \infty\}$; N_1 is a closed subgroup of N so the above integral may be written

$$\xi^{-1}(\gamma)\int_{K\times M\times N_1\backslash N}\{\int_{-\infty}^{\infty}\psi_\Lambda(k\exp(tX)nm\gamma m^{-1}k^{-1})\,dt\}\,dk\,dm\,d\bar{n}.$$

To show that $\chi(\gamma) = 0$ it is sufficient to show that the inner integral is identically zero; this will be done using Cauchy's integral theorem. Recall that

$\psi_\Lambda(k\exp(tX)nm\gamma m^{-1}k^{-1}) = (\zeta(k)\zeta(\exp(tX))\zeta(nm\gamma m^{-1}k^{-1})\phi_0, \phi_0)\,e^{\lambda(\Gamma(g))}$ with $g = \exp(tX)nm\gamma m^{-1}$. The first term is clearly an entire function of t. $\Gamma(g) = -\pi(\mu(m\gamma m^{-1}n^{-1}, \exp(-tX)(0))) - \pi(\mu(\exp(-tX), 0))$. For m, γ, and n fixed the first term is defined, bounded, and analytic in t so long as $\exp(-tX)(0)$ is in B. If it is observed that the sub group of G_c whose Lie algebra is spanned by H_{γ_l}, X_{γ_l}, and $X_{-\gamma_l}$ is the homomorphic image of $SL(2,\boldsymbol{C})$ then the calculations may be performed in this group. Now

$$X = \begin{pmatrix} -i & i \\ -i & i \end{pmatrix} \qquad \exp(tX) = \begin{pmatrix} 1-ti & ti \\ -ti & 1+ti \end{pmatrix}$$

and

$$\begin{pmatrix} 1+ti & -ti \\ ti & 1-ti \end{pmatrix} = \begin{pmatrix} 1 & -it(1-ti)^{-1} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} (1-ti)^{-1} & 0 \\ 0 & (1-ti) \end{pmatrix}\begin{pmatrix} 1 & 0 \\ it(1-ti)^{-1} & 1 \end{pmatrix}$$

Thus $\exp(-tX)(0) = -it(1-it)^{-1}X_{\gamma_l}$ is in B if $|it(1-ti)^{-1}| < 1$ or $\operatorname{Im}(t) > -\frac{1}{2}$ and $-\pi(\mu(\exp(-tX),0)) = \log(1-ti)\pi(\exp(H_{\gamma_l}))$ is analytic in the half-plane $\operatorname{Im}(t) > -\frac{1}{2}$. Moreover, in this region

$$\begin{aligned}|\psi_\Lambda(k\exp(tX)nm\gamma m^{-1}k^{-1})| &\leqq c(1+|t|)^n\,|1-ti|^{\lambda(H_{\gamma_l})}\\ &\leqq c(1+|t|)^{-2}\end{aligned}$$

if $\lambda(H_{\gamma_l}) \ll 0$. n is a positive integer. Here and in what follows c is used as a generic symbol for a positive constant. Cauchy's integral theorem may now be applied.

Suppose γ is singular. γ belongs to the centralizer of at least one Cartan subgroup $\mathfrak{j}$ of $\mathfrak{g}$; $\mathfrak{j}$ may be taken so that $\theta(\mathfrak{j}) = \mathfrak{j}$. Let $\mathfrak{g}_\gamma$ be the centralizer of γ in $\mathfrak{g}$, then $\theta(\mathfrak{g}_\gamma) = \mathfrak{g}_\gamma$. Consequently $\mathfrak{g}_\gamma$ is the direct sum of an abelian algebra $\mathfrak{a}$ and a semi-simple algebra $\mathfrak{g}_1$. Let $\mathfrak{j}_1$ be a fundamental Cartan subalgebra of $\mathfrak{g}_1$ ([2(g)], p. 236). Then $\mathfrak{j}_1 + \mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$ which may be supposed to equal $\mathfrak{j}$. Let B^0 be the connected component of the centralizer of $\mathfrak{j}$ in G. Then

$$\int_{G/B^0} f(g^*)\,dg^* = \int_{G/G_\gamma^0} ds_\gamma^0 \int_{G_\gamma^0/B^0} f((gg_0)^*)\,dg_0^*$$
$$= \int_{G/G_\gamma^0} ds_\gamma^0 \int_{G_1/B_1} f((gg_1)^*)\,dg_1^*$$

if G_1 is the connected group with Lie algebra $\mathfrak{g}_1$ and $B_1 = B^0 \cap G_1$. The measures are so normalized that $dg = dg^* db$ and $dg_\gamma^0 = dg_0^* db$, db being the Haar measure on B^0. Moreover G_γ^0 is the homomorphic image of $G_1 \times A$ where A is the connected group with Lie algebra $\mathfrak{a}$. The measure on G_γ^0 is so normalized that $\int_{G_\gamma^0} f(g_\gamma^0)\,dg_\gamma^0 = \int_{G_1\times A} f(g_1 a)\,dg_1 da$. Finally let B_1^0 be the connected component of the identity in B_1 and write the above integral as

$$[B_1 : B_1^0]^{-1} \int_{G/G_\gamma^0} ds_\gamma^0 \int_{G_1^0/B_1^0} f((gg_1)^*)\,dg_1^*.$$

Choose γ_1 close to the identity in B so that $\gamma\gamma_1$ is regular then

$$(4)\qquad [B_1 : B_1^0] \int_{G/B^0} f(g\gamma\gamma_1 g^{-1})\,dg^* = \int_{G/G_\gamma^0} ds_\gamma^0 \int_{G_1/B_1^0} f(g\gamma g_1\gamma_1 g_1^{-1} g^{-1})\,dg_1^*.$$

If G_1 is any connected semi-simple group with finite center, B_1 the centralizer of a Cartan subalgebra $\mathfrak{j}_1$ of $\mathfrak{g}_1$, and $m(g_1)$ a function on $\tilde{G}_1$, the universal covering group of G_1, a function can be defined on $\tilde{B}_1$ by

$$\phi_m(\gamma_1) = \Delta_1(\gamma_1) \int_{G_1^0/B_1^0} m(g_1\gamma_1 g_1^{-1})\,dg_1^*$$

when the integral exists. To obtain $\Delta_1(\gamma_1)$ map $\tilde{G}_1$ into the simply connected complex group whose Lie algebra is the complexification of $\mathfrak{g}_1$; let γ_1 go into γ_1 and set $\gamma_1 = \exp(H_1)$ with H_1 in the complexification of $\mathfrak{j}_1$. Then $\Delta_1(\gamma_1) = e^{-\rho_1(H_1)} \prod_{\alpha\in P} (e^{\alpha(H_1)} - 1)$; P_1 is the set of positive roots with respect to some order on $\mathfrak{j}_1$ and ρ_1 is one-half the sum of the roots in P_1. For every $\alpha \in P_1$, H_α defines an invariant differential operator D_α on $\tilde{B}_1$; set $D_1 = \prod_{\alpha\in P} D_\alpha$. Harish-Chandra [2(h)] has shown that if $m(g_1)$ is infinitely differentiable with compact support then

$$\lim_{\gamma_1\to 1} D_1\phi_m(\gamma_1) = am(1).$$

a is a constant independent of m and $a \neq 0$ if $\mathfrak{j}_1$ is fundamental. To be more precise $\phi_m(\gamma_1)$ is defined if γ_1 is regular and the limit is taken on the set of regular elements.

Apply this result formally to equation (4) with $f(g) = d_\Lambda\psi_\Lambda(g)$ and

$m(g_1) = d_\Lambda \psi_\Lambda (g\gamma g_1 g^{-1})$. If B^0 is not compact the left side is 0 and one obtains

$$\int_{G/G_\gamma^0} d_\Lambda \psi_\Lambda (g\gamma g^{-1})\, ds_\gamma{}^0 = 0$$

so that $\chi(\gamma) = 0$ if γ has no fixed point in B. If B^0 is compact it may be supposed that $\mathfrak{j} = \mathfrak{h}$ and that $\mathfrak{j}_1 = \mathfrak{j} \cap \mathfrak{g}_1$. In this case B_1 is connected. If $\gamma\gamma_1 = \exp(H + H_1)$ then

$$a \int_{G/G_\gamma^0} d_\Lambda \psi_\Lambda (g\gamma g^{-1})\, ds_\gamma{}^0$$
$$= \lim_{H_1 \to 0} D_1 \{ (\prod_{\substack{\alpha \in P \\ \alpha \notin P_1}} (e^{\frac{1}{2}\alpha(H+H_1)} - e^{-\frac{1}{2}\alpha(H+H_1)}))^{-1} \sum_{s \in w} \epsilon(s)\, e^{s\Lambda(H+H_1)+s\rho(H+H_1)} \},$$

if the total measure of B^0 is 1 as will be assumed and if an order on $\mathfrak{j}_1$ is so chosen that the positive roots are just the positive roots of $\mathfrak{j}$ whose root vectors lie in $\mathfrak{g}_{1,c}$. The denominator is regular at $H_1 = 0$ and is invariant under the Weyl group of $\mathfrak{g}_{1,c}$. Thus, as on page 159 of [2(h)] the right side equals

$$(\prod_{\substack{\alpha \in P \\ \alpha \notin P}} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)}))^{-1} \lim_{H_1 \to 0} D_1 \{ \sum_s \epsilon(s)\, e^{s\Lambda(H+H_1)+s\rho(H+H_1)} \}$$

The second term equals

$$\omega_1 \sum_{w_1/w} \epsilon(s) \prod_{\alpha \in P_1} (s\Lambda(H_\alpha) + s\rho(H_\alpha))\, e^{s\Lambda(H)+s\rho(H)};$$

w_1 is the Weyl group of $\mathfrak{k}_c \cap \mathfrak{g}_{1,c}$ and ω_1 is its order. The sum is over a set of representatives of cosets of w_1 in w. It remains to calculate a. Since, as is easily seen, every non-compact positive root of $\mathfrak{g}_{1,c}$ is totally positive G_1 is locally isomorphic to the product of a compact group and the group of pseudo-conformal mappings of a bounded symmetric domain B_γ. Λ_1 and $\psi_{\Lambda_1}(g_1)$ may be defined in the same manner as Λ and $\psi_\Lambda(g)$; the compact component causes no difficulty [2(c)]. Apply the limit formula of Harish-Chandra to $d_{\Lambda_1}\psi_{\Lambda_1}(g_1)$ to obtain

$$a d_{\Lambda_1} = \lim_{H_1 \to 0} D_1 \{ \sum_{s \in w_1} \epsilon(s)\, e^{s\Lambda_1(H_1)+s\rho_1(H_1)} \}$$
$$= \omega_1 \prod_{\alpha \in P_1} (\Lambda_1(H_\alpha) + \rho_1(H_\alpha))$$

if the total measures of $B_1{}^0$ is 1 as may be assumed. If the measure on G_1 is normalized in the same way as that on G then

$$d_{\Lambda_1} = (-1)^{b_\gamma}/v(B_\gamma) \prod_{\alpha \in P_1} (\Lambda_1(H_\alpha) + \rho_1(H_\alpha)/\rho_1(H_\alpha))$$

if b_γ is equal to the complex dimension of B_γ. a is now easily determined; setting $P_1 = P_\gamma$, $\rho_1 = \rho_\gamma$, and $w_1 = w_\gamma$ the value of (3) is found to be

$$(-1)^{b_\gamma}/v(B_\gamma)\{[G_\gamma\colon G_\gamma^0]^{-1}\prod_{\alpha\in P_\gamma}\rho_\gamma(H_\alpha)\prod_{\substack{\alpha\in P\\ \alpha\notin P_\gamma}}(e^{\frac{1}{2}\alpha(H)}-e^{-\frac{1}{2}\alpha(H)})\}^{-1}$$
$$\times\sum_{w_\gamma/w}\epsilon(s)\prod_{\alpha\in P_\gamma}(s\Lambda(H_\alpha)+s\rho(H_\alpha))e^{s\Lambda(H)+s\rho(H)}$$

and (2) is established. It should be observed that since the total measure of both B^0 and B_1^0 must be 1 the measure on A, and thus on G_γ^0, is completely determined.

6. The prime task of this section is to justify the above application of the limit formula of Harish-Chandra. The truth of the other unproved statements above will become evident in the course of the justification so there is no need to mention them explicitly again. G will now denote a connected semi-simple group with finite centre and $\tilde{G}$ will denote its universal covering group. All the other standard symbols will also refer to G.

In particular $\mathfrak{a}_\mathfrak{p}$ will be a maximal abelian subalgebra of $\mathfrak{p}$. If $\{H_i\}$ is a basis for $\mathfrak{a}_\mathfrak{p}$ and if $g = k_1\exp(H)k_2$ with $H = \sum_{i=1}^{s} t_iH_i$ set $t_i = t_i(g)$. $(t_i(g))$ is not uniquely determined by g. Let $\omega(k)$, for $k\in K$, be the determinant of the restriction of $I - Ad(k)$ to $\mathfrak{p}$; then

LEMMA 3. *There are positive constants ϵ, c, and q so that*

$$\exp(\sum_{i=1}^{s}|t_i(gkg^{-1})|)\geqq c\,|\omega(k)|^q\exp(\epsilon\sum_{i=1}^{s}|t_i(g)|).$$

If $A = (A_{ij})$ is any matrix set $\|A\| = (\sum_{i,j}|A_{ij}|^2)^{\frac{1}{2}}$ and if $g\in G$ let $\|g\| = \|Ad(g)\|$ where $Ad(g)$ is the matrix of the adjoint of g with respect to a basis of $\mathfrak{g}$ orthonormal with respect to the inner product $-B(X,\theta(Y))$. It is easy to verify that

$$c_1\exp(\beta_1\sum_{i=1}^{s}|t_i(g)|)\geqq\|g\|\geqq c_2\exp(\beta_2\sum_{i=1}^{s}|t_i(g)|)\tag{5}$$

for some positive constants c_1, c_2, β_1, and β_2. Now it is sufficient to verify the lemma for $g = a = \exp(H)$. If $Ad(k) = (s_{ij})$ with respect to a basis which diagonalizes $\mathfrak{a}_\mathfrak{p}$ then $\|aka^{-1}\|^2 = \sum_{i,j}e^{\lambda_{ij}(H)}s_{ij}^2$. The λ_{ij} are linear functions on $\mathfrak{a}_\mathfrak{p}$. For a fixed k with $\omega(k)\neq 0$ this must approach infinity with $\sum_{i=1}^{s}|t_i|$ ([2(h)], p. 743). Let $S(H)$ be the set of pairs (ij) for which $\lambda_{ij}(H) > 0$ then

$$\sum_{(ij)\in S(H)}s_{ij}^2\neq 0$$

unless $\omega(k)=0$. Moreover, for some small positive number $\epsilon(H)$, $(ij)\in S(H)$ implies $\lambda_{ij}(H)>3\epsilon(H)\sum_{i=1}^{s}|t_i|$. Let $M=\{H=\sum_{i=1}^{s}t_iH_i\mid\sum_{i=1}^{s}|t_i|=1\}$. If H is in M there is a neighbourhood $U(H)$ of H in M so that if H' is in $U(H)$ and $t>0$ then $\|\exp(tH')k\exp(-tH')\|^2\geqq\exp(\epsilon(H)t\sum_{i=1}^{s}|t_i'|)\{\sum_{(ij)\in S(H)}s_{ij}^2\}$. Since $\sum_{(ij)\in S(H)}s_{ij}^2$ vanishes only when $\omega(k)$ vanishes the theorem of Łojasiewicz [4] implies that there are positive constants $c(H)$ and $q(H)$ so that

$$(\sum_{(ij)\in S(H)}s_{ij}^2)^{\frac{1}{2}}\geqq c(H)\,|\,\omega(k)\,|^{q(H)}$$

for all k. All that is left is to observe that M is compact.

Suppose γ is a semi-simple element of $\bar{G}$. Define G_γ, $G_\gamma{}^0$, G_1, and so on, as before; however it is no longer necessary to suppose that $\mathfrak{j}_1$ is fundamental. Then

$$[B_1\colon B_1{}^0]\int_{G/B^0}f(g^*)\,dg^*=\int_{G/G_\gamma{}^0}ds_\gamma{}^0\int_{G_1/B_1{}^0}f((gg_1)^*)\,dg_1{}^*.$$

If $\Delta_1(\gamma_1)$ and D_1 are defined as above we are to show that

$$\lim_{\gamma_1\to1}D_1\Delta_1(\gamma_1)\int_{G/G_\gamma{}^0}ds_\gamma{}^0\int_{G_1/B_1{}^0}\psi(g\gamma g_1\gamma_1g_1{}^{-1}g^{-1})\,dg_1{}^*$$
$$=a\int_{G/G_\gamma{}^0}\psi(g\gamma g^{-1})\,ds_\gamma{}^0. \tag{6}$$

γ_1 is chosen so that $\gamma\gamma_1$ is reguar and the limit is taken in the manner previously indicated. Of course it will be necessary to impose some conditions on the function ψ. If ψ is infinitely differentiable with compact support then for γ_1 in some compact neighbourhood of the identity the inner integral on the left, a function on $G/G_\gamma{}^0$, vanishes outside some fixed compact set U ([2(h)], Thm. 1). Moreover

$$D_1\Delta_1(\gamma_1)\int_{G_1/B_1{}^0}\psi(g\gamma g_1\gamma_1g_1{}^{-1}g^{-1})\,dg_1{}^*$$

converges uniformly on U to $a\psi(g\gamma g^{-1})$ ([2(h)], Thms. 2 and 4). This shows the validity of (6) for functions with compact support. To establish it for another function ψ it would be sufficient to show that for any $\epsilon>0$ there is a sequence $\{\psi_i(g)\}$ of infinitely differentiable functions with compact support such that

$$\text{(i)}\qquad\lim_{i\to\infty}D_1\Delta_1(\gamma_1)\int_{G/B^0}\psi_i(g\gamma\gamma_1g^{-1})\,dg^*=D_1\Delta_1(\gamma_1)\int_{G/B^0}\psi(g\gamma\gamma_1g^{-1})\,dg^*$$

uniformly in γ_1 and

$$\text{(ii)} \qquad \lim_{i\to\infty}\int_{G/G_\gamma^0}\psi_i(g\gamma g^{-1})\,ds_\gamma{}^0=\int_{G/G_\gamma^0}\psi(g\gamma g^{-1})\,ds_\gamma{}^0.$$

γ_1 is, of course, to lie in a fixed compact neighbourhood of the identity and be such that $\gamma\gamma_1$ is regular.

In order to establish the existence of $\{\psi_i\}$ it is sufficient to assume that ψ is infinitely differentiable and that there is a sufficiently large constant α such that, for any left-invariant differential operator D on $\bar{G}$, $|D\psi(g)| \leqq c(D)\|g\|^{-\alpha}$.

Once it has been verified that this condition is satisfied by $\psi_\Lambda(g)$ when λ is real and $\lambda(H_{\gamma_i}) \ll 0$ there will no longer be any need to refer specifically to this function. For convenience, if $X\in\mathfrak{g}$ we denote the differential operator $\frac{d}{dt}f(g\exp(tX))|_{t=0}$ by X. It may be supposed that $D=\prod_{i=1}^k X_i$ so that

$$D\psi_\Lambda(g)=D\{(\zeta(g)\phi_0,\phi_0)e^{\lambda(\Gamma(g))}\}=\sum_\sigma D_{\sigma'}(\zeta(g)\phi_0,\phi_0)D_\sigma e^{\lambda(\Gamma(g))};$$

σ runs over the subsets of $\{1,\cdots,k\}$ and $D_\sigma=\prod_{i\in\sigma}X_i$ with the order of the X_i's left unchanged. σ' is the complement of σ. Now $D_{\sigma'}(\zeta(g)\phi_0,\phi_0) = (\zeta(g)\prod_{i\in\sigma'}\zeta(X_i)\phi_0,\phi_0)$ so there is no doubt that

$$|D_{\sigma'}(\zeta(g)\phi_0,\phi_0)| \leqq c(D_\sigma')\|g\|^{\alpha_1}$$

with some constant α_1. To find $D_\sigma e^{\lambda(\Gamma(g))}$ we must differentiate

$$\exp(\lambda(\Gamma(g\prod_{i\in\sigma}\exp(t_iX_i))))$$

with respect to each of the variables and evaluate the result at the origin. But $\Gamma(g\prod_{i\in\sigma}\exp(t_iX_i))=-\pi(\mu(\prod'_{i\in\sigma}\exp(-t_iX_i)g^{-1},0))$, the prime indicating that the order of the factors is reversed, and this equals

$$-\pi(\mu(g^{-1},0))-\pi(\mu(\prod_{i\in\sigma}{}'\exp(-t_iX_i),g^{-1}(0)).$$

There is an open neighbourhood U of the identity in G_c and an open neighbourhood V of $\bar{B}$, the closure of B, such that $h\in U$ implies $h^{-1}(V)\subseteq\mathfrak{p}_+$. Consequently $\mu(h^{-1},z)$ is defined and analytic on $U\times V$. So is $\pi(\mu(h^{-1},z))$ and its derivatives at the identity are bounded functions on B. Thus

$$|D_\sigma e^{\lambda(\Gamma(g))}|\leqq c(D_\sigma)|e^{\lambda(\Gamma(g))}|;$$

but $|e^{\lambda(\Gamma(g))}|\leqq c(\lambda)\exp(\sum_{i=1}^s|t_i(g)|\lambda(H_{\gamma_i}))$ ([2(d)], p. 600). Thus, if $\lambda(H_{\gamma_i})\ll 0$, $i=1,\cdots,s$,

$$|D_\sigma e^{\lambda(\Gamma(g))}| \leqq c(D_\sigma)\exp(-\alpha_2 \sum_{i=1}^{s} |t_i(g)|)$$

with α_2 large and positive. These remarks and formula (5) show that the assumption is satisfied.

We shall need a non-decreasing sequence $\{\phi_i(g)\}$ of infinitely differentiable functions on $\tilde{G}$ with compact support satisfying conditions: (α) $\lim_{i\to\infty}\phi_i(g) = 1$, ($\beta$) there is a non-decreasing sequence $\{U_i\}$ of open sets which exhausts $\tilde{G}$ such that $\phi_i(g) \equiv 1$ on U_i, (γ) if D is a left invariant differential operator on $\tilde{G}$ then $|D\phi_i(g)| \leqq c(D)$ for all i and g. We write, after Iwasawa, $\tilde{G} = KHN$. If $\{X_i\}$ is a basis for $\mathfrak{g}$ and if (a_{ij}) is the matrix of $Ad(hn)$ with respect to this basis then

$$X_i f(khn) = \sum_j a_{ji} \frac{d}{dt} f(k\exp(tX_j)hn)$$
$$= \sum_j a_{ji} f(khn; Y_j, Z_j)$$

if $X_j = Y_j + Z_j$, $Y_j \in \mathfrak{k}$ and $Z_j \in \mathfrak{h} + \mathfrak{n}$. If D_1 is a left invariant differential on $\tilde{K}$ and D_2 a right invariant differential operator on HN then $f(g; D_1, D_2)$ is the result of the successive applications of these two operators to f considered as a function on $\tilde{K} \times HN$. In particular,

$$f(khn; 1, Z_j) = \frac{d}{dt} f(k\exp(tZ_j)hn)|_{t=0}.$$

Iterating we obtain

$$Df(khn) = \sum_j g_i(hn) f(khn; D_1^i, D_2^i);$$

$g_i(hn)$ is a polynomial in the coefficients of $Ad(hn)$. If $X \in \mathfrak{h} + \mathfrak{n}$ and $X = X_1 + X_2$, $X_1 \in \mathfrak{h}$, $X_2 \in \mathfrak{n}$, then

$$\frac{d}{dt} f(\exp(tX)hn) = \frac{d}{dt} f(\exp(tX_1)hn) + \frac{d}{dt} f(h\exp(th^{-1}(X_2))n).$$

Consequently

$$Df(khn) = \sum_i g_i(hn, h^{-1}) f(khn; D_1^i, D_2^i, D_3^i);$$

D_1^i acts on $\tilde{K}$, D_2^i on H, and D_3^i on N. $g_i(hn, h^{-1})$ is a polynomial in the coefficients of $Ad(hn)$ and $Ad(h^{-1})$. The functions $\phi_i(g)$ are to be constructed as products $\phi_i(khn) = \phi_i^1(k)\phi_i^2(h)\phi_i^3(n)$. Since the coefficients $g_j(hn, h^{-1})$ are independent of k we need only require that $\{\phi_i^1(k)\}$ satisfy (α), (β), and (γ). This requirement is easily satisfied since $\tilde{K}$ is the product

of a vector group and a compact group. H is a vector group and the coefficients $g_i(nh, h^{-1})$ are exponential polynomials on H so we need only require that $\{\phi_i^2(h)\}$ satisfies (α) and (β) and that the derivatives on $\phi_i^2(h)$ go to zero faster than any exponential polynomial uniformly in i. N is a closed subset of the space of endomorphisms of $\mathfrak{g}$ and the functions $g_j(hn, h^{-1})$ are polynomials in the coefficients of $Ad(n)$. Thus the functions $\phi_i^3(n)$ can be obtained as the restriction to N of a sequence of functions on a vector group which satisfies (α) and (β) and such that the derivatives of the functions go to zero faster than any inverse polynomial uniformly in i.

Now set $\psi_i(g) = \phi_i(g)\psi(g)$. To establish (i) it is sufficient to show that for any invariant differential operator v on $\bar{B}$

$$\lim_{i\to\infty} v\Delta(\gamma)\int_{G/B^0}\psi_i(g\gamma g^{-1})\,dg^* = v\Delta(\gamma)\int_{G/B^0}\psi(g\gamma g^{-1})\,dg^* \tag{7}$$

uniformly in γ on any bounded subset (i.e. a subset with compact closure in $\bar{B}$) of the set of regular elements in $\bar{B}$. $\Delta(\gamma)$ is defined in the same manner as $\Delta_1(\gamma_1)$. To obtain (i) from this relation it is sufficient to set $v = D_1$ and to observe that $\Delta_1(\gamma_1)\Delta^{-1}(\gamma\gamma_1)$ is regular at $\gamma_1 = 1$.

If M and N are the groups introduced on p. 212 of [2(g)] then

$$\Delta(\gamma)\int_{G/B^0}\psi(g\gamma g^{-1})\,dg^* = \Delta(\gamma)\xi(X_i)^{-1}(\gamma)\int_{K\times M\times N}\psi(knm\gamma m^{-1}k^{-1})\,dk\,dm\,dn$$

Let S be a finite set of invariant differential operators on B and let l be the maximum degree of the operators in S. Let δ belong to S and let D be a left invariant differential operator on $\bar{G}$. δ determines in an obvious fashion a left invariant differential operator on $\bar{G}$ which will be denoted δ'. Then

$$\begin{aligned}|\delta(D\psi(knm\gamma m^{-1}k^{-1}))| &= |Ad(km)(\delta')D\psi(knm\gamma m^{-1}k^{-1})| \\ &\leqq \|km\|^{l}c(D)\|nm\gamma m^{-1}\|^{-\alpha}\end{aligned}$$

and

$$|\delta(D\psi(knm\gamma m^{-1}k^{-1})) - \delta(D\psi_i(knm\gamma m^{-1}k^{-1}))| \leqq \|km\|^{l}c(D)\|nm\gamma m^{-1}\|^{-\alpha}$$

Moreover there is an increasing sequence $\{V_i\}$ of open sets in $N\times M$, which exhaust $N\times M$, so that the left-side of the latter inequality is zero if $(n, m)\in V_i$.

Recall that B^0 is the connected component of the identity of the centralizer in G of a Cartan subalgebra $\mathfrak{j}$ and that $\theta(\mathfrak{j}) = \mathfrak{j}$. An examination of the form of the matrices of $Ad(m\gamma m^{-1})$ and $Ad(n)$ with respect to a basis which diagonalizes $\mathfrak{j}\cap\mathfrak{p}$ shows that $\|nm\gamma m^{-1}\| \geqq \|m\gamma m^{-1}\|$. Thus $\|nm\gamma m^{-1}\|^{-\alpha}$

$\leqq \| m\gamma m^{-1} \|^{-\alpha_1} \| nm\gamma m^{-1} \|^{-\alpha_2}$ if $\alpha_1 + \alpha_2 = \alpha$. Now γ may be written as $\gamma_-\gamma_+$ with $\gamma_- \in \tilde{B} \cap \tilde{K}$ and $\gamma_+ \in \exp(\mathfrak{j} \cap \mathfrak{p})$. Then

$$\| m\gamma m^{-1} \| \geqq \| \gamma_+^{-1} \|^{-1} \| m\gamma_- m^{-1} \| \geqq c \,|\, \omega_-(\gamma_-) \,|^q \| \gamma_+^{-1} \|^{-1} \| m \|^{\epsilon};$$

here $\omega_-(\gamma_-)$ is the determinant of the restriction of $I - Ad(\gamma_-)$ to $\mathfrak{m} \cap \mathfrak{p}$ and c, q, and ϵ are positive constants. Thus

$$\| km \|^{l} \| nm\gamma m^{-1} \|^{-\alpha} \leqq c \,|\, \omega_-(\gamma_-) \,|^{q(\alpha_2-\alpha_1)} \| \gamma_+^{-1} \|^{\alpha_1-\alpha_2} \| m \|^{\epsilon(\alpha_2-\alpha_1)+l} \| n \|^{-\alpha_2};$$

consequently

$$(8) \qquad \int_{K\times M\times N} |\, \delta(D\psi(knm\gamma m^{-1}k^{-1}))\,|\, dkdmdn$$
$$\leqq c(D) \,|\, \omega_-(\gamma_-) \,|^{q(\alpha_2-\alpha_1)} \| \gamma_+^{-1} \|^{\alpha_1-\alpha_2} \int_M \| m \|^{\epsilon(\alpha_2-\alpha_1)+l}\, dm \int_N \| n \|^{-\alpha_2}\, dn$$

and

$$(9) \qquad \int_{K\times M\times N} |\, \delta(D\psi(knm\gamma m^{-1}k^{-1})) - \delta(D\psi_i(knm\gamma m^{-1}k^{-1}))\,|\, dkdmdn$$
$$\leqq c(D) \,|\, \omega_-(\gamma_-) \,|^{q(\alpha_2-\alpha_1)} \| \gamma_+^{-1} \|^{\alpha_1-\alpha_2} \int_{V_i'} \| m \|^{\epsilon(\alpha_2-\alpha_1)+l} \| n \|^{-\alpha_2}\, dmdn.$$

Now it can be shown (cf. [2(g)], Cor. 1 to Lemma 6) that the integral over N in (8) converges if α_2 is sufficiently large. Then, fixing α_2, we can choose α_1 so large that the first integral converges. Moreover, by the dominated convergence theorem, the integrals in (8) converge to zero as i approaches infinity. We conclude first of all that

$$\phi_\psi(\gamma) = \Delta(\gamma) \int_{G/B^0} \psi(g\gamma g^{-1})\, dg^*$$

is defined on $\tilde{B}' = \{\gamma \in \tilde{B} \mid \omega_-(\gamma_-) \neq 0\}$ and is the uniform limit on compact subsets of $\tilde{B}'$ of the sequence $\{\phi_{\psi_i}(\gamma)\}$.

There is a finite set $\{v_1, \cdots, v_w\}$ of invariant differential operators on $\tilde{B}$ so that any other v, may be written as $v = \sum_{j=1}^{w} v_j u_j$ where the u_j are invariant under the Weyl group of $\mathfrak{g}_c$ ([2(f)], p. 101). For each u_j there is a left and and right invariant differential operator D_j on $\tilde{G}$ so that $u_j\phi_{\psi_i}(\gamma) = \phi_{D_j\psi_i}(\gamma)$ ([2(h)], p. 155). Then

$$v\phi_{\psi_i}(\gamma) = \sum_{j=1}^{w} v_j \phi_{D_j\psi_i}(\gamma).$$

The right side is a sum of terms of the form

$$\phi(\gamma) \int_{K\times M\times N} \delta(D\psi_i(knm\gamma m^{-1}k^{-1}))\, dkdmdn.$$

$\phi(\gamma)$ is a regular function on $\bar{B}$; δ is one of a finite set of invariant differential operators on $\bar{B}$; and D is a left and right invariant differential operator on $\bar{G}$. As a consequence of the estimates above the sequence $\{v\phi_{\psi_i}(\gamma)\}$ converges uniformly on compact subsets of $\bar{B}'$ and it must converge to $v\phi_\psi(\gamma)$ so

$$|v\phi_\psi(\gamma) - v\phi_{\psi_i}(\gamma)| \leqq |\omega_-(\gamma_-)|^{q(\alpha_2-\alpha_1)}c(v,i)$$

on any fixed bounded subset of $\bar{B}'$. Moreover $\lim_{i\to\infty} c(v,i) = 0$. The proof of (7) can now be completed by an argument essentially the same as that on pp. 208-211 of [2(g)]. There is no point in reproducing it. If we show that

$$\int_{G/G_\gamma^0}\psi(g\gamma g^{-1})\,ds_\gamma^0 \tag{10}$$

is absolutely convergent then a simple application of the dominated convergence theorem suffices to establish (ii). Choose a maximal abelian subspace of $\mathfrak{g}_1 \cap \mathfrak{p}$ and extend it to a Cartan subalgebra $\mathfrak{j}_1$ of $\mathfrak{g}_1$, then $\mathfrak{j} = \mathfrak{j}_1 + \mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$ and $\theta(\mathfrak{j}) = \mathfrak{j}$. We again introduce the groups M and N. If $\mathfrak{n}$ is the Lie algebra of N let $\mathfrak{n}_1 = \mathfrak{n} \cap \mathfrak{g}_1$ and let N_1 be the connected group with Lie algebra $\mathfrak{n}_1$. If $\mathfrak{n}_2 = (I - Ad(\gamma))\mathfrak{n}$ then, according to Lemma 7.0 of [1] every element of N may be written uniquely in the form $\exp(Y_2)n_1$ with $n_1 \in N_1$ and $Y_2 \in \mathfrak{n}_2$. Then if B_+ is the connected group with Lie algebra $\mathfrak{j} \cap \mathfrak{p}$ every element of G may be written uniquely as $g = k\exp(X)\exp(Y_2)n_1b_+$ with $X \in \mathfrak{m} \cap \mathfrak{p}$, $\mathfrak{m}$ being the Lie algebra of M, and $b_+ \in B_+$ (cf. [2(h)], p. 215). Let $\phi_0(n_1b_+)$ be a non-negative function on N_1B_+ such that

$$\int_{N_1B_+}\phi_0(n_1b_+)\,dn_1db_+ = 1,$$

and set $\phi(g) = \|\exp(X)\exp(Y_2)\|^\beta\phi_0(n_1b_+)$ if $g = k\exp(X)\exp(Y_2)n_1b_+$. β is a suitably chosen non-negative constant. Then a function may be defined on G/G_γ^0 by

$$\phi(s_\gamma^0) = \int_{G_\gamma^0}\phi(gg_\gamma)\,dg_\gamma;$$

s_γ^0 is the coset containing g. We shall show that if β is sufficiently large then $\phi(s_\gamma^0)$ is greater than some fixed positive constant for all s_γ^0. It may be assumed that $g = k\exp(X)\exp(Y_2)$. If K_γ is the connected group with Lie algebra $\mathfrak{k} \cap \mathfrak{g}_\gamma$ and if $gu = k'\exp(X')\exp(Y_2')n_1'b_+'$, $u \in K_\gamma$, then

$$\phi(s_\gamma{}^0) = \int_{K_\gamma\times N_1\times B_+} \phi(gun_1b_+)\,du\,dn_1\,db_+$$
$$= \int_{K_\gamma} \|\exp(X')\exp(Y_2')\|^\beta e^{-2\rho_1(\log b_+')}\,du.$$

Suppose $f_1(z_1)$ and $f_2(z_2)$ are two non-negative functions of the variables z_1 and z_2 and z_1 and z_2 are subject to some relation. We shall write $f_1 \succ f_2$ if there is a positive constant c and a non-negative constant β so that $cf_1{}^\beta(z_1) \geqq f_2(z_2)$ for all pairs (z_1, z_2) satisfying the given relation. The assertion will be proved if it is shown that $\|\exp(X')\exp(Y_2')\| \succ e^{2\rho_1(\log b_+')}$. If $H_+ \in \mathfrak{j} \cap \mathfrak{p}$ then $2\rho_1(H^+)$ is the trace of the restriction of $Ad(H_+)$ to $\mathfrak{n}_1$.

Since $\|\exp(X)\exp(Y)\| = \|\exp(X')\exp(Y_2')n_1'b_+'\|$ it is easily seen, after choosing a basis of $\mathfrak{g}$ which diagonalizes $\mathfrak{j} \cap \mathfrak{p}$, that $\|\exp(X)\exp(Y)\| \geqq \|\exp(X')b_+'\|$. If $\mathfrak{a}_-$ is a maximal abelian subalgebra of $\mathfrak{m} \cap \mathfrak{p}$ then $\mathfrak{a}_- + (\mathfrak{j} \cap \mathfrak{p})$ is a maximum abelian subalgebra $\mathfrak{a}_\mathfrak{p}$ of $\mathfrak{p}$. Let $X' = k_-(H_-)$ with $H_- \in \mathfrak{a}_-$ and $k_- \in M \cap K$ and let $b_+' = \exp(H_+)$ with $H_+ \in \mathfrak{j} \cap \mathfrak{p}$. If α is the restriction of a root to $\mathfrak{a}_\mathfrak{p}$ then

$$\log \|\exp(X)b_+'\| \geqq |\alpha(H_- + H_+)|;$$

since the restrictions of the roots to $\mathfrak{a}_\mathfrak{p}$ span the space of linear functions on $\mathfrak{a}_\mathfrak{p}$ there is a constant c so that for any linear function λ

$$c\|\lambda\| \log \|\exp(X')b_+'\| \geqq |\lambda(H_- + H_+)|.$$

Since $\mathfrak{a}_- \cap (\mathfrak{j} \cap \mathfrak{p}) = \{0\}$ it is now clear that $\|\exp(X)\exp(Y_2)\| \succ \|b_+'\|$ and $\|\exp(X)\exp(Y_2)\| \succ \|\exp(X')\|$. From this one easily deduces that $\|\exp(X)\exp(Y_2)\| \succ \|\exp(Y_2')n_1'\|$. If $n_1' = \exp(Y_1')$ and $\exp(Y_2')\exp(Y_1') = \exp(Y')$ with $Y_1' \in \mathfrak{n}_1$ and $Y' \in \mathfrak{n}$ then the four variables (Y_1', Y_2'), Y', $(Ad(\exp(Y_1'))$, $Ad(\exp(Y_2'))$ and $Ad(\exp(Y'))$ are polynomial functions of each other (cf. [2(h)], pp. 737-738 and the reference cited there.) Consequently $\|\exp(Y')\| \succ \|\exp(Y_1')\|$ and $\|\exp(Y')\| \succ \|\exp(Y_2')\|$ so that $\|\exp(X)\exp(Y_2)\| \succ \|\exp(Y_1')\|$ and $\|\exp(X)\exp(Y_2)\| \succ \|\exp(Y_2')\|$ However if $n_1'b_+'u^{-1} = u'b_+^{-1}n_1^{-1}$ then

$$k\exp(X)\exp(Y_2)n_1b_+ = k'\exp(X')\exp(Y_2')u'$$

and the argument may be reversed. Consequently $\|\exp(X')\exp(Y')\| \succ \|b_+'\| \succ e^{2\rho_1(\log b_+')}$.

The absolute convergence of (10) will be established if it is shown that

$$\int_G |\psi(g\gamma g^{-1})|\,\phi(g)\,dg$$

converges. However this integral equals

$$\int_{K\times M\times N\times B_+} |\psi(kmn\gamma n^{-1}m^{-1}k^{-1})|\,\phi(kmnb_+)\,dk\,dm\,dn\,db_+.$$

If $n=\exp(Y_2)n_1$ as before then $dn=dY_2dn_1$ where dY_2 is the Euclidean measure on $\mathfrak{n}_2$ and this integral equals

$$\int_{K\times M\times\mathfrak{n}_2} |\psi(km\exp(Y_2)\gamma\exp(-Y_2)m^{-1}k^{-1})|\,\|m\exp(Y_2)\|^{\beta}\,dk\,dm\,dY_2.$$

The integrand is less than or equal to

$$c\,\|(m\exp(Y_2)\gamma\exp(-Y_2)\gamma^{-1}m^{-1})m\gamma m^{-1}\|^{-\alpha}\,\|m\|^{\beta}\,\|\exp(Y_2)\|^{\beta}$$

$$\leqq c\,\|\exp(Y_2)\gamma\exp(-Y_2)\gamma^{-1}\|^{-\alpha_1}\,\|m\|^{2\alpha_1+\beta}\,\|m\gamma m^{-1}\|^{-\alpha_2}\,\|\exp(Y_2)\|^{\beta}$$

if $\alpha_1+\alpha_2=\alpha$, $\alpha_1,\alpha_2\geqq 0$. It follows from Lemma 8 of [2(h)] that $\|m\gamma m^{-1}\|\succ\|m\|$ and from Lemma 2 of that paper that

$$\|\exp(Y_2)\gamma\exp(-Y_2)\gamma^{-1}\|\succ 1+\|Y_2\|.$$

Thus if α is sufficiently large the integrand is less than or equal to a multiple of $(1+\|Y_2\|)^{-\beta_1}\|m\|^{-\beta_2}$ with β_1 and β_2 large. Consequently the integral converges. It should be observed that γ is fixed so that uniform estimates like that of Lemma 3 are not necessary.

PRINCETON UNIVERSITY
AND
THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] A. Borel and Harish-Chandra, *Annals of Mathematics*, vol. 75 (1962), pp. 485-535.

[2] Harish-Chandra, (a) *Transactions of the American Mathematical Society*, vol. 76 (1954), pp. 485-528.

(b) *American Journal of Mathematics*, vol. 77 (1955), pp. 743-777.

(c) *American Journal of Mathematics*, vol. 78 (1956), pp. 1-41.

(d) *American Journal of Mathematics*, vol. 78 (1956), pp. 564-628.

(e) *Transactions of the American Mathematical Society*, vol. 83 (1956), pp. 98-163.

(f) *American Journal of Mathematics*, vol. 79 (1957), pp. 87-120.

(g) *American Journal of Mathematics*, vol. 79 (1957), pp. 193-257.

(h) *American Journal of Mathematics*, vol. 79 (1957), pp. 733-760.

(i) *American Journal of Mathematics*, vol. 80 (1958), pp. 241-310.

[3] F. Hirzebruch, *Symposium Internacional de Topologia Algebraica*, Universidad de Mexico (1958), pp. 129-143.

[4] S. Łojasiewicz, *Studia Mathematica*, vol. 18 (1959), pp. 87-136.

[5] A. Selberg, *Journal of the Indian Mathematical Society*, vol. 20 (1956), pp. 47-87.

[6] ———, *Seminars on Analytic Functions*, Institute for Advanced Study, Princeton, vol. 2, pp. 152-161.

[7] H. Cartan, *Séminaire*, Paris, 1957/1958.

SEPARABLY GENERATED SPOTS AND AFFINE RINGS OVER REGULAR RINGS.*

By Louis J. Ratliff, Jr.

1. **Introduction.** The results in this paper are generalizations of some results in [2] and [5], and the notation and terminology is the same as in [5]. In particular a Noetherian domain I is said to satisfy the *condition* (SF) if each separably generated affine ring $\mathfrak{o}$ over I is such that the integral closure of $\mathfrak{o}$ in its quotient field is a finite $\mathfrak{o}$-module. It is well known ([8], p. 267) that a field satisfies the condition (SF), and recently Nagata ([2], p. 419) has proven that a Dedekind domain satisfies this condition. Nagata's proof is essentially based on the following facts.

1.1. Let $\mathfrak{o}$ be an affine ring over I, and let $\mathfrak{o}^*$ be an integral domain which contains $\mathfrak{o}$ and is a finite $\mathfrak{o}$-module. If $\mathfrak{p}^*$ is a prime ideal in $\mathfrak{o}^*$, then $\operatorname{rank} \mathfrak{p}^* = \operatorname{rank}(\mathfrak{p}^* \cap \mathfrak{o})$.

1.2. If P is a separably generated spot over I, then the integral closure of P in its quotient field is a finite P-module ([2], p. 416).

It is known ([2], p. 414) that if a local domain P is analytically unramified, then the integral closure of P in its quotient field is a finite P-module. Hence statement 1.2 above is a consequence of

1.3. If P is a spot which is separably generated over I, then P is analytically unramified ([2], p. 416).

In an earlier paper [5] the author proved that a rank two regular local ring I which has an infinite residue field satisfies the condition (SF). The proof was somewhat complicated by the existence of spots of type one, and the condition of an infinite residue field for I was necessitated to handle this case. In the current paper no classification of spots into type one or type two is needed. By proving a preliminary theorem (Theorem 3.1) on analytical unramifiedness of spots dominating regular local rings, the dichotomy of spots is avoided in this paper and the restriction on the rank ($\leqq 2$) of a ground ring is also removed.

A regular Noetherian domain I of finite rank (if p is a prime ideal in I,

* Received December 20, 1962.

then I_p is a regular local ring) is a *ground ring*. By using some results in [3, 4] it is proven in Section 2 that if $\mathfrak{o}$ is an affine ring over I and if $\mathfrak{q}$ is a maximal ideal in $\mathfrak{o}$, then $\operatorname{trd}(\mathfrak{o}/I) = \operatorname{rank} q - \operatorname{rank}(q \cap I)$. Corollary 2 of this theorem is the generalization of statement 1.1 above to the case where I is a regular Noetherian domain. In Section 3 it is proven that statement 1.3 above holds when I is a regular Noetherian domain of finite rank, and hence Nagata's proof that a Dedekind domain has the property (SF) carries over verbatim to prove that a regular Noetherian domain of finite rank has the property (SF).

2. The rank of an affine ring over a ground ring. As stated in Section 1 the statement that I is a ground ring is to mean that I is a rank n $(0 \leqq n < \infty)$ regular Noetherian domain.

Nagata [3, 4] has proven two theorems which are stated as Remark 2.1 for future reference.

Remark 2.1. The second chain condition holds in a Noetherian domain J if and only if the first chain condition holds in the derived normal ring of J. Further, if J satisfies the second chain condition, and if P is a spot over J, then P satisfies the second chain condition and the dimension formula.

LEMMA 2.1. *Let I be a ground ring, and let $\mathfrak{p}$ be a prime ideal in I. Then $I_{\mathfrak{p}}$ satisfies the dimension formula and the second chain condition.*

Proof. It is known ([7], p. 187) that a regular local ring $I_{\mathfrak{p}}$ satisfies the first chain condition. Since $I_{\mathfrak{p}}$ is a normal Noetherian domain ([9], p. 302), $I_{\mathfrak{p}}$ satisfies the second chain condition (Remark 2.1). Since $I_{\mathfrak{p}}$ is a spot over $I_{\mathfrak{p}}$, $I_{\mathfrak{p}}$ satisfies the dimension formula (Remark 2.1), q. e. d.

Remark 2.2. Since $I = \bigcap_M I_M$, where M is a maximal ideal in I, and since I_M is a normal domain, I is a normal domain.

The proofs of the following theorem and corollaries are the same as given in [5], hence are omitted.

THEOREM 2.1. *Let $\mathfrak{o} = I[z_1, \cdots, z_k]$ be an affine ring over a ground ring I, let $\mathfrak{q}$ be a maximal ideal in $\mathfrak{o}$, and let $\mathfrak{q} \cap I = \mathfrak{p}$. Then* $\operatorname{trd}(\mathfrak{o}/I) = \operatorname{rank} \mathfrak{q} - \operatorname{rank} \mathfrak{p}$.

COROLLARY 2.1. *With the same notation as in Theorem 2.1, if $(0) = \mathfrak{q}_0 \subset \mathfrak{q}_1 \subset \cdots \subset \mathfrak{q}_s = \mathfrak{q}$ is a maximal chain of prime ideals in $\mathfrak{o}$, then* $s = \operatorname{rank} \mathfrak{q}$.

COROLLARY 2.2. *Let* $\mathfrak{o}$ *be an affine ring over a ground ring* I, *let* $\mathfrak{o}^*$ *be an integral domain which contains* $\mathfrak{o}$ *and is a finite* $\mathfrak{o}$ *module, and let* $\mathfrak{p}^*$ *be a prime ideal in* $\mathfrak{o}^*$. *Then* $\operatorname{rank} \mathfrak{p}^* = \operatorname{rank} (\mathfrak{p}^* \cap \mathfrak{o})$.

COROLLARY 2.3. *Let* (P, M) *be a spot over a ground ring* I, *let* P' *be a finite* P-*module, and let* $\mathfrak{p}$ *be a prime ideal in* P'. *Then* $\operatorname{rank} \mathfrak{p} = \operatorname{rank}(\mathfrak{p} \cap P)$.

3. Analytical unramifiedness of spots which are separably generated over a ground ring. Statement 1.3 of the introduction will be proven in this section (Theorem 3.2).

THEOREM 3.1. *Let* (P, M) *be a spot over a ground ring* C. *If the quotient field* L *of* P *is a finite separable algebraic extension of the quotient field* E *of* C, *then* P *is analytically unramified.*

Proof. (Note that by the dimension formula for the regular local ring $C_{M \cap C}$, the hypotheses imply $\operatorname{rank} P \leqq \operatorname{rank} C_{(M \cap C)}$). Since P is a spot over C, P is a spot over $C_{(M \cap C)}$, so we may consider C as being a regular local ring dominated by P, and $\operatorname{rank} P \leqq \operatorname{rank} C$. Let C' be the integral closure of C in L, and let $P' = P[C']$. Since C is a normal Noetherian domain and L is separable over E, C' is a finite C module, so P' is a finite P-module, and P is a subspace of P' ([9], p. 277). Further, P' is a semi-local domain, so it is only necessary to prove that $P'_{M'}$ is analytically unramified, where M' is a maximal ideal of P' ([9], p. 283). Now $\operatorname{rank} M' = \operatorname{rank} M$ (Corollary 2.3), and

$$M' \cap C = (M' \cap P) \cap C = M \cap C, \qquad M' \cap C = (M' \cap C') \cap C = M \cap C,$$

so $\operatorname{rank}(M' \cap C') \cap C = \operatorname{rank}(M \cap C)$ (Corollary 2.2). Let $P'' = P'_{M'}$, and let $D = C'_{(M' \cap C')}$, so $P'' \geqq D$, $\operatorname{rank} P'' \leqq \operatorname{rank} D$, and P'' is a spot over D (since P'' is a spot over C). Hence there exist elements $e_1, \cdots, e_k$ in P'' such that P'' is a quotient ring with respect to a prime ideal of the ring $D[e_1, \cdots, e_k]$, so by a result of Rees [6] P'' is analytically unramified if D is. Since C' has only a finite number of maximal ideals and $M' \cap C'$ is one of them, it is sufficient to prove that C' is analytically unramified ([9], p. 283).

Since L is a finite separable algebraic extension of E, let u be an element in C' such that $L = \sum_{i=0}^{a} E u^i$, and set $H = \sum_{i=0}^{a} C u^i$. Then H is an integral domain which contains C and is contained in C', hence H is a semi-local domain with quotient field L. Let F be the total quotient ring of the completion of H. Since C' is a finite C-module, there exists a nonzero element $d \in C$ such that $dC' \subseteq H$. Since the completion of C is a regular local ring

([9], p. 302), d is not a zero divisor in this completion of H ([9], p. 277), hence F contains the completion of C'. Let K be the quotient field of the completion of C. Then it is seen that $F = \sum_{i=0}^{a} Ku^i$. Since F is a finite dimensional algebra over K which contains the completion of C' and is contained in the total quotient ring of the completion of C', F is the total quotient ring of the completion of C'. It is a straight forward verification that F is isomorphic to $L \otimes_E K$, and since L is separably generated over E, $L \otimes_E K$ contains no nonzero nilpotent elements ([8], p. 195). Therefore C' is analytically unramified, hence P is analytically unramified, q. e. d.

THEOREM 3.2. *If (P, M) is a spot which is separably generated over a ground ring I, then P is analytically unramified.*

Proof. Let $z_1, -, z_t$ be elements in P which form a separating transcendence base for the quotient field of P over the quotient field of I, and let $C = I[z_1, -, z_t]$. Then C is a ground ring (if $\mathfrak{p}$ is a prime ideal in C, then ${}^{I}(\mathfrak{p} \cap I)$ is a regular local ring, hence $C_{\mathfrak{p}}$ is a regular local ring ([3], p. 404)), P is a spot over C, and the quotient field of P is a finite separable algebraic extension of the quotient field of C, hence P is analytically unramified (Theorem 3.1), q. e. d.

It is known ([2], p. 414) that if a semi-local domain P is analytically unramified, then the derived normal ring of P is a finite P-module, hence

COROLLARY 3.1. *If P is a spot which is separably generated over a ground ring I, then the derived normal ring of P is a finite P-module.*

COROLLARY 3.2. *Let C be a regular local ring with quotient field E, and let L be a finite separable extension field of E. Then the integral closure C' of C in L is analytically unramified.*

Proof. The completion of C' is a finite direct sum of analytically unramified rings, q. e. d.

COROLLARY 3.3. *Let P be a separably generated spot over a ground ring I. Then every quadratic transform P' of P is analytically unramified.*

Proof. P' is a spot over P and P is a spot over I, so P' is a separably generated spot over I, q. e. d.

COROLLARY 3.4. *Let P be a separably generated spot over a ground ring I, let K be a finite separable extension of the quotient field of P, and*

let $x_1, -, x_t$ *be elements in* K. *Then* $P[x_1, -, x_t]_{\mathfrak{q}}$ *is analytically unramified, for all prime ideals* $\mathfrak{q}$ *in* $P[x_1, -, x_t]$.

Proof. $P[x_1, -, x_t]_{\mathfrak{q}}$ is a separably generated spot over I.

As mentioned in the introduction, Nagata's proof that a Dedekind domain has the property (SF) is essentially based on statements 1.1 and 1.2. Since Corollary 2.2 and Corollary 3.1 are respectively the generalizations of these statements to the case where I is a regular Noetherian domain of finite rank, Nagata's proof carries over verbatim, since the rank of I is finite, to prove

THEOREM 3.3. *If* $\mathfrak{o}$ *is an affine ring which is separably generated over a ground ring* I, *then the derived normal ring of* $\mathfrak{o}$ *is a finite* $\mathfrak{o}$-*module.*

INDIANA UNIVERSITY.

REFERENCES.

[1] M. Nagata, "A general theory of algebraic geometry over Dedekind domains, I," *American Journal of Mathematics*, vol. 78 (1956), pp. 78-116.

[2] ———, "A general theory of algebraic geometry over Dedekind domains, II," *American Journal of Mathematics*, vol. 80 (1958), pp. 382-420.

[3] ———, "On the chain problem of prime ideals," *Nagoya Mathematical Journal*, vol. 10 (1956), pp. 51-64.

[4] ———, "Note on a chain condition for prime ideals," *Memoirs of the College of Science, University of Kyoto*, Ser. A, vol. 32 (1959), pp. 85-90.

[5] L. J. Ratliff, Jr., "Affine rings over rank two regular local rings," *American Journal of Mathematics*, vol. 83 (1961), pp. 709-722.

[6] D. Rees, "A note on analytical unramified local rings," *Journal of the London Mathematical Society*, vol. 36 (1961), pp. 24-28.

[7] J. P. Serre, "Sur la dimension homologique des anneaux et des modules noethériens," *Proceedings of International Symposium on Algebraic Number Theory*, Tokyo-Nikko 1955, 1956.

[8] O. Zariski and P. Samuel, *Commutative Algebra*, Volume I. D. van Nostrand Co., Inc., 1958.

[9] O. Zariski and P. Samuel, *Commutative Algebra*, Volume II. D. van Nostrand Co., Inc., 1960.

EXTENSIONS OF REPRESENTATIONS OF ALGEBRAIC LINEAR GROUPS.*[1]

By A. Bialynicki-Birula, G. Hochschild and G. D. Mostow.

1. **Introduction.** Let G be an algebraic linear group over an arbitrary field F. If ρ is a rational representation of G by linear automorphisms of a finite-dimensional F-space U we refer to this structure (U, ρ) by saying that U is a finite-dimensional rational G-module. A G-module that is a sum (not necessarily direct) of finite-dimensional rational G-modules is called a *rational G-module.* Let K be an algebraic subgroup of G. We are interested in determining when every finite-dimensional rational representation of K can be extended to a rational representation of G, i.e., when every finite-dimensional rational K-module can be imbedded as a K-submodule in a rational G-module. In the analogous situation for Lie groups, an analysis of the the obstructions to the extendibility of representations of a subgroup has been made only for normal subgroups, [5], [7], and not much is known in the general case. The algebraic case turns out to be much more accessible. In particular we shall show, with a very simple argument, that *the extendibility of arbitrary finite-dimensional representations is already implied by the extendibility of the 1-dimensional representations.* From this, we shall obtain a variety of other sufficient conditions, the simplest of which amounts to the result that *if K is a normal algebraic subgroup of G then every finite-dimensional rational representation of K can be extended to one of G.* If the base field is algebraically closed, we shall see that *the extendibility of representations of a subgroup K of G is equivalent to the condition that the algebraic variety G/K be quasi-affine.*

One may also ask the dual question: is every finite-dimensional rational K-module a K-homomorphic image of a finite-dimensional rational G-module? By considering the dual representations, this is immediately seen to be equivalent to the extension question, so that our results imply the appropriate dual results.

2. **Representative functions and reduction to rational characters.** We begin by recalling some elementary notions and results of representation

* Received December 18, 1962.
[1] Research done under NSF Grant G-24943.

theory. The *rational representative functions* on the algebraic group G are the composites $\phi \circ \rho$, where ρ is a finite-dimensional rational representation of G and ϕ is a linear functional on the space of all linear endomorphisms of the representation space of ρ. These functions $\phi \circ \rho$ are called the representative functions *associated with the representation* ρ. We denote the space of all these by $R(\rho)$. The group G operates by left and right translations on $R(\rho)$. If $f \in R(\rho)$ and $x \in G$, the left translate $x \cdot f$ is defined by $(x \cdot f)(y) = f(yx)$, and the right translate $f \cdot x$ is defined by $(f \cdot x)(y) = f(xy)$. The totality of all representative functions on G constitute an algebra $R(G)$ over the basic field F. $R(G)$ is a rational G-module, with G operating either by left translations or by right translations.

If G is given concretely as an algebraic subgroup of the full linear group on a finite-dimensional F-space V, and if ρ is the identity representation of G, then $R(G)$ is generated, as an F-algebra, by the elements of $R(\rho)$, the constant 1, and the reciprocal of the determinant function [6, Lemma 10.1]. It is immediately seen from this that, *if K is an algebraic subgroup of G, the restriction map $R(G) \to R(K)$ is surjective.*

Let ρ be a finite-dimensional rational representation of the algebraic subgroup K of G. It is an elementary fact [6, Prop. 2, 3] that the rational K-module determined by ρ may be identified with a K-submodule of a direct sum of a finite number of copies of the rational K-module $R(\rho)$ (on which K acts by left translations). Using that the restriction map $R(G) \to R(K)$ is surjective, we see from this that the representation space W of ρ may be written as U/V, where U is a finite-dimensional rational K-module that is contained as a K-submodule in a finite-dimensional rational G-module M, and where V is a K-submodule of U.

Let n be the dimension of V, and consider $E^{n+1}(M)$, the homogeneous component of degree $n+1$ of the exterior F-algebra built over M. This contains $U \wedge E^n(V)$ as a K-submodule. Clearly, $U \wedge E^n(V)$ is K-isomorphic with the tensor product $W \otimes E^n(V)$ of the rational K-module W by the 1-dimensional rational K-module $E^n(V)$. Now let S be the 1-dimensional rational K-module whose character is the reciprocal of the character belonging to $E^n(V)$. Then W is K-isomorphic with $(U \wedge E^n(V)) \otimes S$. Hence, if S is contained as a K-submodule in a rational G-module, so is W. Thus we have the following result.

THEOREM 1. *Let G be an algebraic linear group, and let K be an algebraic subgroup of G. Suppose that, for every 1-dimensional rational K-module that is contained as a K-submodule in a rational G-module, the dual*

K-module is also a K-submodule of a rational G-module. Then every finite-dimensional rational K-module is a K-submodule of a finite-dimensional rational G-module.

Since a unipotent algebraic group has no non-trival 1-dimensional rational representations, we have the following corollary.

COROLLARY 1. *Let K be a unipotent algebraic subgroup of the algebraic linear group G. Then every finite-dimensional rational K-module is a K-submodule of a finite-dimensional rational G-module.*

3. Application in the case of characteristic 0. Let us assume that the basic field F is of characteristic 0. Let G be an algebraic linear group over F, and let K be an algebraic subgroup of G. We denote by N_K the maximum unipotent normal subgroup of K. Then, if S_K is any maximal fully reducible subgroup of K, we have a semidirect decomposition $K = S_K \cdot N_K$ [8, Theorems 6.1 and 7.1]. This decomposition is rational, in the following sense. $R(K)$ is isomorphic with the tensor product $R(S_K) \otimes R(N_K)$ by the map associating with $f \otimes g$, where $f \in R(S_K)$ and $g \in R(N_K)$, the function h on K given by $h(xy) = f(x)g(y)$, for all $x \in S_K$ and all $y \in N_K$ [4, Section 3]. The use we shall make of this decomposition is based on the well known fact that, over a field of characteristic 0, every rational module for a fully reducible algebraic linear group is semisimple.

THEOREM 2. *Let G be an algebraic linear group over a field of characteristic 0. Let K be an algebraic subgroup of G, and let N_K, N_G denote the maximum unipotent normal subgroups of K, G, respectively. Let U be the smallest normal algebraic subgroup of G containing N_K. Then each of the following three conditions is sufficient for every finite-dimensional rational K-module to be a K-submodule of a finite-dimensional rational G-module. (1) : $N_K \subset N_G$; (2) : K is normal in G; (3) : $U \cap K = N_K$, and the base field is algebraically closed.*

Proof. In case (1), we make a semidirect product decomposition $K = S_K \cdot N_K$, as described above, and we let S_G denote a maximal fully reducible subgroup of G containing S_K, so that $G = S_G \cdot N_G$. Now let ρ be a 1-dimensional rational representation of K. Consider the restriction of ρ to S_K. Its representation space may be written as A/B, where A is a finite-dimensional rational S_K-module that is contained as an S_K-submodule in a finite-dimensional rational S_G-module, and where B is an S_K-submodule of A. Since the basic field is of characteristic 0 and S_K is fully reducible, A is semisimple as an S_K-module, so that A/B is isomorphic with an S_K-submodule of

A. Thus we conclude that the restriction of ρ to S_K can be extended to a finite-dimensional rational representation σ of S_G. We may regard σ as a rational representation τ of G whose kernel contains N_G. Since $N_K \subset N_G$ and N_K lies in the kernel of ρ, it is clear that τ is an extension of ρ. By Theorem 1, this implies that every finite-dimensional rational K-module is a K-submodule of a finite-dimensional rational G-module.

In case (2), we note that N_K is also normal in G, whence $N_K \subset N_G$, so that the conclusion follows from case (1).

Now suppose that condition (3) is satisfied. By [3, Prop. 11, p. 119], there is a finite-dimensional rational representation ρ of G whose kernel is precisely U. Since the base field is algebraically closed, $\rho(G)$ is an algebraic linear group, and $\rho(K)$ is an algebraic subgroup of $\rho(G)$. Since $U \cap K = N_K$, the restriction of ρ to S_K is a bijection of S_K onto $\rho(K)$. Since the base field is algebraically closed and of characteristic 0, the inverse $\tau\colon \rho(K) \to S_K$ of the restriction of ρ to S_K is also rational [6, Lemma 10.2].

Now let σ be a 1-dimensional rational representation of K. Then σ must be trivial on N_K, and $\sigma \circ \tau$ is a rational representation of $\rho(K)$. Since $\rho(K) = \rho(S_K)$, it is fully reducible. Hence, by case (1) above, $\sigma \circ \tau$ extends to a finite-dimensional rational representation γ of $\rho(G)$. Now $\gamma \circ \rho$ is a rational representation of G that extends the given representation σ of K. This completes the proof of Theorem 2.

4. Algebraic-geometric structure. Let us say that an algebraic subgroup K of an algebraic linear group G is an *observable* subgroup if every finite-dimensional rational K-module is a K-submodule of a finite-dimensional rational G-module. We make a few notational conventions, as follows.

$Q(G)$ will denote the field of all rational functions on G; this is the field of quotients of $R(G)$. For any G-module M, M^G wil denote the subspace of M consisting of all G-fixed elements. We shall regard $R(G)$ and $Q(G)$ as K-modules with K operating by left translations, where K is a subgroup of G.

THEOREM 3. *Let G be an irreducible algebraic linear group over an arbitrary field, and let K be an algebraic subgroup of G. Then K is an observable subgroup of G if and only if $Q(G)^K$ coincides with the field of quotients of $R(G)^K$.*

Proof. Suppose first that K is an observable subgroup of G. Let q be a non-zero element of $Q(G)^K$. It suffices to show that $(R(G)q \cap R(G))^K \neq (0)$. Let V be any non-zero simple K-submodule of $R(G)q \cap R(G)$, and let V^* be the dual K-module. By assumption, V^* is a K-submodule of a finite-

dimensional rational G-module W. Since W is isomorphic with a G-submodule of a direct sum of a finite number of copies of $R(G)$, and since V^* is simple, it follows that there exists a K-module monomorphism ϕ of V^* into $R(G)$. Let $(v_1, \cdots, v_n)$ be a basis of V such that $v_1(1) \neq 0$ and $v_i(1) = 0$, for each $i > 1$. Let $(\alpha_1, \cdots, \alpha_n)$ be the dual basis of V^*, and put $g_i = \phi(\alpha_i)$. Then $\sum_{i=1}^{n} (g_i \cdot x)v_i \in (R(G)q \cap (R(G))^K$, for every $x \in G$. Take x such that $g_1(x) \neq 0$. Then this function is not zero. Thus the necessity of the condition is proved.

In proving the sufficiency, let us first observe that we may assume that the base field F is algebraically closed. For let F' be an algebraic closure of F, and let G', K' be the canonical extensions of G, K, respectively, to algebraic linear groups over F'. We have $R(G') = R(G) \otimes_F F'$ and $Q(G') = Q(G) \otimes_F F'$. Since every rational function vanishing on K vanishes also on K', we have $R(G')^{K'} = R(G')^K = R(G)^K \otimes_F F'$, and similarly $Q(G')^{K'} = Q(G)^K \otimes_F F'$. Hence if $Q(G)^K$ is the field of quotients of $R(G)^K$ then also $Q(G')^{K'}$ is the field of quotients of $R(G')^{K'}$. On the other hand, it is easy to see that if K' is observable in G' then K is observable in G. In fact, if S is a finite-dimensional rational K-module, then S is a K-submodule of the rational K'-module $S \otimes_F F'$. Now, if $S \otimes_F F'$ is a K'-submodule of a rational G'-module T, we may regard T as a G-module over the field F. Since $R(G') = R(G) \otimes_F F'$, T is then a rational G-module containing S as a K-submodule, which suffices for concluding that K is observable in G.

There remains to prove the sufficiency of the condition in the case where F is algebraically closed. In that case, let f be an element of $R(G)$ such that $f(1) \neq 0$ and $x \cdot f = \gamma(x)f$, with $\gamma(x) \in F$, for every $x \in K$. For every $x \in G$, we have $(f \cdot x)/f \in Q(G)^K$, so that, by our assumption, there is a non-zero element k_x in $R(G)^K$ such that $(f \cdot x)k_x \in R(G)^K f$. Let Z be the set of zeros of f in G, and let Z' be its complement. Then k_x must vanish on $Z \cap x^{-1}Z'$. Since $1 \in Z'$, $y \in yZ'$, for every $y \in G$. Hence Z is covered by the open sets $x^{-1}Z'$. Hence there is a finite set $x_1, \cdots, x_n$ in G such that Z is the union of the sets $Z \cap x_i^{-1}Z'$. Put $k = k_{x_1} \cdots k_{x_n}$. Then k is a non-zero element of $R(G)^K$ that vanishes on Z. By Hilbert's Nullstellensatz, there is therefore a positive integer m such that $k^m = gf$, with $g \in R(G)$. This yields $x \cdot g = \gamma(x)^{-1}g$, for every $x \in K$. The existence of a non-zero element $g \in R(G)$ such as we have just found, for every $f \in R(G)$ such as the above, is evidently equivalent to the condition of Theorem 1. Hence we may indeed conclude that K is an observable subgroup of G.

THEOREM 4. *Let G be an irreducible algebraic linear group over an*

algebraically closed field F, and let K be an algebraic subgroup of G. Then the following conditions are equivalent:

(1) *K is an observable subgroup of G;*

(2) *$R(G)^K$ separates the points of G/K;*

(3) *the field of quotients of $R(G)^K$ coincides with $Q(G)^K$;*

(4) *the algebraic variety G/K is quasi-affine (i.e., an open subvariety of an affine algebraic variety).*

Proof. By Theorem 3, (1) and (3) are equivalent. Since $Q(G)^K$ separates the points of G/K [1, p. 8], (3) implies (2).

The rest of the proof requires more elaborate algebraic-geometric tools. We recall first that the canonical structure of an algebraic variety of G/K satisfies the following (characteristic) conditions. The canonical map $G \to G/K$ is an open morphism whose cohomomorphism maps the field of the rational functions on G/K isomorphically onto $Q(G)^K$. If ϕ is any morphism of G into an algebraic variety V such that $\phi(xy) = \phi(x)$, for all $x \in G$ and all $y \in K$, there is a unique morphism ψ of G/K into V whose composite with the canonical morphism $G \to G/K$ coincides with ϕ [1, Theorems 1 and 4].

Now we show that (4) implies (1). Suppose that G/K is a quasi-affine variety, and let f be an element of $R(G)$ such that $f(1) \neq 0$ and $x \cdot f = \gamma(x) f$, for every $x \in K$, with $\gamma(x) \in F$. Let Z be the set of zeros of f in G. Let p denote the canonical map $G \to G/K$. Since $ZK = Z$, we have $Z = p^{-1}(p(Z))$. Since p is an open map, it follows that $p(Z)$ is closed in G/K. Since $K \notin p(Z)$ and G/K is quasi-affine, there is an everywhere defined rational function α on G/K such that $\alpha(K) \neq 0$ and $\alpha(p(Z)) = (0)$. Put $k = \alpha \circ p$. Then $k \in R(G)^K$, $k(Z) = (0)$ and $k(1) \neq 0$. As we have seen in the proof of Theorem 3, this implies that K is an observable subgroup of G, i.e., that (1) is satisfied.

Next we show that (2) implies (3). Let T denote the field of quotients of $R(G)^K$. If (2) is satisfied T separates the points of the algebraic variety G/K. It follows that the field $Q(G)^K$ of all rational functions on G/K is a purely inseparable algebraic extension of T. This implies that, if J is the integral closure of $R(G)^K$ in $Q(G)^K$, T coincides with the field of quotients of J. On the other hand, G is a *normal* affine algebraic variety, and $R(G)$ is the ring of all everywhere defined rational functions on G. Hence $R(G)$ is integrally closed in $Q(G)$, whence $J \subset R(G) \cap Q(G)^K = R(G)^K$, and $T = Q(G)^K$. Thus we have shown that (2) implies (3).

There remains only to show that (3) implies (4). If (3) is satisfied then, since $Q(G)^K$ is finitely generated, there is a finitely generated F-subalgebra B

of $R(G)^K$ such that $Q(G)^K$ coincides with the field of quatients of B and B is stable under the right translations $f \to f \cdot x$, with $x \in G$. Let V denote the affine algebraic variety whose points are the specializations of B. Clearly, the group G acts as a group of automorphisms $v \to x \cdot v$ on V, where $(x \cdot v)(b) = v(b \cdot x)$, for every $b \in B$. Let ϕ denote the morphism $G \to V$ defined by $\phi(x)(b) = b(x)$, for every $x \in G$ and every $b \in B$. Evidently, ϕ is a G-morphism with respect to the action of G by the translations $y \to xy$ on G and $v \to x \cdot v$ on V. By elementary specialization theory, $\phi(G)$ contains an open subset of V. Since G acts transitively on $\phi(G)$, it follows that $\phi(G)$ is open in V. Thus $\phi(G)$ is a quasi-affine algebraic variety. Since $\phi(xy) = \phi(x)$, for every $x \in G$ and every $y \in K$, ϕ induces a unique morphism $\psi\colon G/K \to \phi(G)$ such that $\psi \circ p = \phi$, where p is the canonical morphism $G \to G/K$. Evidently, ψ is still a G-morphism with respect to the action of G on G/K and on $\phi(G)$, and ψ is bijective. Since the field of quotients of B coincides with $Q(G)^K$ and since $\phi(G)$ is open in V, the cohomorphism of ψ is an isomorphism of the field of rational functions on $\phi(G)$ onto the field of the rational functions on G/K. Hence there is a rational map ρ from $\phi(G)$ to G/K such $\psi \circ \rho$ and $\rho \circ \psi$ are the identity maps on $\phi(G)$ and G/K, respectively. Since ψ is bijective, this means that the set-theoretical inverse ψ^{-1} of ψ coincides with a morphism on some open subvariety of $\phi(G)$. Since ψ is a G-morphism, it is clear that if ψ^{-1} coincides with a morphism on an open subvariety D of $\phi(G)$ then ψ^{-1} coincides with a morphism also on $x \cdot D$, for every $x \in G$. Since G acts transitively on $\phi(G)$, it follows that ψ^{-1} is actually a morphism of $\phi(G)$ onto G/K. Thus ψ is an isomorphism, and we have shown G/K is quasi-affine. This completes the proof of Theorem 4.

5. Applications. Since, in the last section, we have assumed that the basic field is algebraically closed and G irreducible, we begin with two elementary results of reduction to this case.

THEOREM 5. *Let G be an algebraic linear group over an arbitrary field F, and let K be an algebraic subgroup of G. Let F' be an algebraic extension field of F, and let G', K' denote the canonical extensions of G, K to algebraic linear groups over F'. Then K' is observable in G' if and only if K is observable in G.*

Proof. The proof is contained in our proof of Theorem 3 above.

THEOREM 6. *Let G be an algebraic linear group over an arbitrary field F, let G_1 be the irreducible component of the identity in G, and let K be an algebraic subgroup of G. Then K is observable in G if and only if $K \cap G_1$ is observable in G_1.*

Proof. For any group U, let $F[U]$ denote the group algebra of U over F, and regard any U-module over F as an $F[U]$-module in the usual fashion. Suppose first that K is observable in G, and let W be a finite-dimensional rational $(K \cap G_1)$-module. Form the tensor product $F[K] \otimes_{F[K \cap G_1]} W$, and let K operate on it by multiplication from the left on the factor $F[K]$. Since $K/(K \cap G_1)$ is finite, it is clear that $F[K] \otimes_{F[K \cap G_1]} W$ is a finite-dimensional rational K-module containing W as a $(K \cap G_1)$-submodule in the natural fashion. Since K is observable in G, this rational K-module is a K-submodule of a finite-dimensional rational G-module T, say. We may regard T as a rational G_1-module, and W is a $(K \cap G_1)$-submodule of T. Thus $K \cap G_1$ is observable in G_1.

Now suppose that $K \cap G_1$ is observable in G_1, and let S be a finite-dimensional rational K-module. Then there is a $(K \cap G_1)$-module monomorphism ϕ of S into a finite-dimensional rational G_1-module P. As above, we form the finite-dimensional rational G-module $F[G] \otimes_{F[G_1]} P$. Let $x_1, \cdots, x_n$ be a system of representatives in K for the elements of KG_1/G_1. Define the linear map ψ of S into $F[G] \otimes_{F[G_1]} P$ by $\psi(s) = \sum_{i=1}^{n} x_i \otimes \phi(x_i^{-1} \cdot s)$. Since ϕ is a $(K \cap G_1)$-module homomorphism, it is clear that ψ is actually independent of the particular choice of the representatives x_i. Using this, one sees immediately that ψ is a K-module homomorphism. Moreover, it is clear that ψ is injective. Thus ψ is a K-module imbedding of S in a finite-dimensional rational G-module, and we have shown that K is observable in G. This completes the proof of Theorem 6.

THEOREM 7. *Let G be an algebraic linear group over an arbitrary field, let K be an algebraic subgroup of G, and let L be a normal algebraic subgroup of K. Suppose that L is observable in G and that either K/L is finite or every rational representation of K that is trivial on L is unipotent. Then K is observable in G.*

Proof. By Theorem 6, it suffices to show that $K \cap G_1$ is observable in G_1. Now $L \cap G_1$ is a normal algebraic subgroup of $K \cap G_1$ and is observable in G_1, by Theorem 6. If K/L is finite, so is $(K \cap G_1)/(L \cap G_1)$. A rational representation of $K \cap G_1$ that is trivial on $L \cap G_1$ may be regarded as a rational representation of $L(K \cap G_1)$ that is trivial on L. The induced representation of K, constructed as in the proof of Theorem 6, is then trivial on L. Hence, if every rational representation of K that is trivial on L is unipotent, then also every rational representation of $K \cap G_1$ that is trivial on $L \cap G_1$ is unipotent. Hence we may now suppose that G is irreducible.

In that case, we see from Theorem 3 that it suffices to show that $Q(G)^K$ coincides with the field of quotients of $R(G)^K$. Since L is observable in G, we know from Theorem 3 that $Q(G)^L$ coincides with the field of quotients of $R(G)^L$. Now let q be a non-zero element of $Q(G)^K$. Then, since q lies in the field of quotients of $R(G)^L$, we have $R(G)^L q \cap R(G)^L \neq (0)$. This space is a rational K-module, with K operating by left translation, and L acts trivially. Hence, if every rational representation of K that is trivial on L is unipotent, there is a non-zero K-fixed element in $R(G)^L q \cap R(G)^L$, which means that q lies in the field of quotients of $R(G)^K$. On the other hand, if K/L is finite, write $q = f/g$, with f and g in $R(G)^L$. Let g^* denote the product of the translates $x \cdot g$, where x ranges over a complete system of representatives in K for the cosets of L in K. Then $g^* \in R(G)^K$, and we may write $g^* = gg'$, with $g' \in R(G)^L$. Now we have $q = (fg')/g^*$ and, since q and g^* lie in $Q(G)^K$, it follows that $fg' \in R(G)^K$, so that again q lies in the field of quotients of $R(G)^K$. This completes the proof of Theorem 7.

Corollary 2. *Let G be an algebraic linear group over an arbitrary field, and let K be an algebraic subgroup of G such that K_1 is nilpotent. Then K is observable in G.*

Proof. Theorem 7, with $L = K_1$, shows that it suffices to prove that K_1 is observable in G. By Theorem 5, we may assume that the base field is algebraically closed. By a well known result due to A. Borel, K_1 is then a direct product $L \times U$, where L is central and consists of the semisimple elements of K_1, and U is the maximum unipotent normal subgroup of K_1. Since every rational representation of L is semisimple, it follows, as in the first part of the proof of Theorem 2, that L is observable in G. Since U is unipotent, every rational representation of K_1 that is trivial on L is unipotent. Hence we may conclude from Theorem 7 that K_1 is observable in G, q. e. d.

The implication $(1) \rightarrow (4)$ of Theorem 4, together with Corollary 2 and the case $L = K_1$ of Theorem 7, generalize a result of M. Rosenlicht [9, Th. 3].

The following is essentially a reformulation of parts of Theorems 3 and 4.

Theorem 8. *Let G be an algebraic linear group over the field F, and let K be an algebraic subgroup of G. Then, if K is observable in G, K is the isotropy subgroup of an element of a rational G-module. Conversely, if F is algebraically closed and K is the isotropy subgroup of an element of a rational G-module then K is observable in G.*

Proof. Suppose that K is observable in G. By Theorem 6, $K \cap G_1$ is then observable in G_1. Since $Q(G_1)^{K \cap G_1}$ is finitely generated and separates

the elements of $G_1/(K \cap G_1)$, it is clear from Theorem 3 that there is a finite set $f_1, \cdots, f_n$ of elements of $R(G_1)^{K \cap G_1}$ separating $K \cap G_1$ from all the other cosets of $K \cap G_1$ in G_1. Let V_i be the rational G_1-module spanned by the left translates of f_i, and let V be the direct sum of the V_i's. Then, with $v = (f_1, \cdots, f_n) \in V$, $K \cap G_1$ is clearly the isotropy subgroup of v. Now let $W = F[G] \otimes_{F[G_1]} V$, and let $x_1, \cdots, x_q$ be a complete system of representatives in K for the cosets of $K \cap G_1$. Put

$$w = \sum_{i=1}^{q} x_i \otimes v \in W.$$

Clearly, $x \cdot w = w$, for every $x \in K$. Conversely, if $x \in G$ and $x \cdot w = w$, we note first that no two of the elements $xx_1, \cdots, xx_q$ can lie in the same coset of G_1 in G and hence conclude that $xx_i \in KG_1$, for each i. Thus we have $xx_i = x_{i'}x(i)$, with $x(i) \in G_1$, where $i \rightarrow i'$ is a permutation of the set $(1, \cdots, q)$. Furthermore, we see that we must have $x(i) \cdot v = v$, so that $x(i) \in K \cap G_1$, whence $x \in K$. Thus K is precisely the isotropy subgroup of w, and the first part of Theorem 8 is proved.

Now suppose that F is algebraically closed, that V is a rational G-module, and that K is the isotropy subgroup of an element $v \in V$. For every linear functional ϕ on V, denote by ϕ/v the function on G defined by $(\phi/v)(x) = \phi(x \cdot v)$. Then each ϕ/v is an element of $R(G)^K$, and the set of all these functions, as ϕ ranges over the space of the linear functionals on V, separates the elements of G/K. Thus $R(G)^K$ separates the elements of G/K. A fortiori, $R(G_1)^{K \cap G_1}$ separates the elements of $G_1/(K \cap G_1)$. Hence we conclude from Theorem 4 that $K \cap G_1$ is observable in G_1. By Theorem 6, it follows that K is observable in G, and Theorem 8 is proved.

The criterion of Theorem 8 is convenient in proving the following stronger form of Theorem 7. The reason for exhibiting Theorem 7 separately is that its proof is elementary, while our proof of Theorem 9 below finally appeals to Theorem 4 and thus rests on the consideration of homogeneous spaces as algebraic varieties.

THEOREM 9. *Let G be an algebraic linear group over an arbitrary field F, K an algebraic subgroup of G, L a normal algebraic subgroup of K. Suppose that L is observable in G and that, for every 1-dimensional rational representation ρ of K that is trivial on L and extendible to a rational representation of G, some tensor power of the dual of ρ can be extended to a rational representation of G. Then K is observable in G.*

Proof. We shall show in an elementary way that the assumptions of the theorem imply that K is the isotropy subgroup of an element of a rational

G-module. However, in order to be in position to appeal to Theorem 8 for the conclusion that K is observable in G, we must have F algebraically closed. Therefore, we begin by showing that our assumptions remain satisfied when the base field is extended to its algebraic closure.

Let F' be an algebraic closure of F, and let G', K', L' denote the canonical extensions of G, K, L, respectively, to algebraic linear groups over F'. Then L' is normal in K', and, by Theorem 5, L' is observable in G'. Let γ be a rational character of K' that is trivial on L' and whose associated 1-dimensional K'-module is a K'-submodule of a rational G'-module. We must show that the same is true for the 1-dimensional K'-module associated with some positive integral power of the reciprocal of γ.

Let α be any F-automorphism of F'. Then α induces F-automorphisms $f \to f^\alpha$ of $R(G') = R(G) \otimes_F F'$ and of $R(K') = R(K) \otimes_F F'$ in the natural fashion, and these are compatible with the restriction map from G' to K'. Now γ^α is an element of $R(K')$ whose restriction to K is a homomorphism into the multiplicative group of F' and whose restriction to L is trivial. It follows that γ^α is a rational character of K' that is trivial on L'. Moreover, it is easy to see that the 1-dimensional K'-module associated with γ^α is a K'-submodule of a rational G'-module, from the fact that this is true for γ.

Since the restriction of γ to K takes its values in some finite algebraic extension of F, we can find a finite number of F-automorphisms $\alpha_1, \cdots, \alpha_t$ of F' such that the restriction to K of the product $\gamma\gamma^{\alpha_1} \cdots \gamma^{\alpha_t}$ takes its values in a finite purely inseparable algebraic extension of F. Hence the restriction to K of some positive integral power of this product is a rational character ψ of K that is trivial on L and whose associated 1-dimensional K-module is a K-submodule of a rational G-module.

By the assumption of our theorem, there is a positive integer q such that the 1-dimensional K-module associated with the character $(\psi^{-1})^q$ is a K-submodule of a rational G-module. On the other hand, some positive integral power of $(\gamma\gamma^{\alpha_1} \cdots \gamma^{\alpha_t})^{-1}$ is the canonical extension to K' of $(\psi^{-1})^q$, whence its associated K'-module is a K'-submodule of a rational G'-module. Writing

$$(\gamma^{-1})^r = ((\gamma\gamma^{\alpha_1} \cdots \gamma^{\alpha_t})^{-1})^r (\gamma^{\alpha_1} \cdots \gamma^{\alpha_t})^r$$

and noting that each γ^{α_i} has the requisite extendibility property, we see that, for some positive integer r, the 1-dimensional K'-module associated with $(\gamma^{-1})^r$ is a K'-submodule of a rational G'-module.

Thus (G', K', L') satisfies the conditions of our theorem. Once the theorem is proved for the case of an algebraically closed base field, we conclude that K' is observable in G', and this implies, by Theorem 5, that K is

observable in G. Thus we may now assume that F is algebraically closed, but this will be used only at the very end, when we appeal to Theorem 8.

Since L is observable in G, we know from the first part of Theorem 8 that it is the isotropy subgroup of an element of a rational G-module. As we have seen in proving the second part of Theorem 8, this implies that $R(G)^L$ separates the elements of G/L. It is easily seen from this that there is a finite-dimensional left K-stable subspace U of $R(G)^L$ such that the restriction to K of every element of U is 0 and such that every zero of U lies in K. Let V be the smallest left G-stable subspace of $R(G)$ containing U, and let n be the dimension of U. Consider $E^n(V)$, the homogeneous component of degree n of the exterior algebra built over V. This is a rational G-module containing $E^n(U)$ as a 1-dimensional K-submodule. Clearly, L operates trivially on $E^n(U)$. If $x \in G$ and $x \cdot E^n(U) \subset E^n(U)$ it follows easily that $x \cdot U \subset U$, whence x is a zero of U and so $x \in K$. In fact, we have actually $x \cdot E^n(U) = E^n(U)$, whence $x^{-1} \cdot E^n(U) = E^n(U)$, so that we obtain, working in $E^{n+1}(V)$,

$$(x \cdot U) \wedge E^n(U) = x \cdot (U \wedge x^{-1} \cdot E^n(U)) = x \cdot (U \wedge E^n(U)) = (0),$$

which shows that $x \cdot U \subset U$.

Now apply the assumption of the theorem to the representation ρ of K on $E^n(U)$. Let W be a rational G-module containing the q-th tensor power, X, of the dual of ρ as a K-submodule. With T^q denoting q-th tensor power, consider the rational G-module $S = T^q(E^n(V)) \otimes W$. This contains the trivial 1-dimensional K-module $T^q(E^n(U)) \otimes X$ as a K-submodule. If x is an element of G leaving this submodule fixed we must have $x \cdot E^n(U) \subset E^n(U)$, so that $x \in K$. Thus K is the isotropy subgroup of a basis element of $T^q(E^n(U)) \otimes X$ in the rational G-module S. By Theorem 8, K is therefore observable in G, which completes the proof of Theorem 9.

THEOREM 10. *Let G be an algebraic linear group over an arbitrary field F, and let K be a normal algebraic subgroup of G. Then K is observable in G.*

Proof. By Theorems 5 and 6, we may assume without loss of generality that F is algebraically closed and that G is irreducible. Since K is normal in G, there is a finite-dimensional rational representation of G whose kernel is exactly K [3, Prop. 11, p. 119]. Hence condition (2) of Theorem 4 is satisfied, so that K is observable in G.

THEOREM 11. *Let G be an algebraic linear group over an infinite perfect field. Then, if K and L are observable algebraic subgroups of G, so is $K \cap L$.*

Proof. First, suppose that the base field is algebraically closed. By

Theorem 6, we may assume that G is irreducible. Clearly, if K and L satisfy condition (2) of Theorem 4, so does $K \cap L$, whence the result.

Now assume only that the base field F is infinite and perfect. Let F' be an algebraic closure of F. For any algebraic linear group U over F, let U' denote its canonical extension to an algebraic linear group over F'. By Theorem 5, K' and L' are observable in G'. By what we have just proved, $K' \cap L'$ is therefore observable in G'. Evidently (or by Theorem 6), $(K' \cap L')_1$ is observable in $K' \cap L'$. Hence $(K' \cap L')_1$ is observable also in G'. Now $K' \cap L'$ is defined over the subfield F of F', from which it follows easily that $(K' \cap L')_1$ is also defined over F. By a result due to M. Rosenlicht [10. p. 44, Corollary], this and the fact that F is an infinite perfect field imply that the set of points of $(K' \cap L')_1$ that are rational over F is dense in $(K' \cap L')_1$. Since these points lie in $K \cap L$, we have therefore $(K' \cap L')_1 \subset (K \cap L)'$, whence $(K' \cap L')_1 = ((K \cap L)')_1$. Thus we conclude that $((K \cap L)')_1$ is observable in G'. By Theorem 7, this implies that $(K \cap L)'$ is observable in G'. Finally, Theorem 5 now gives that $K \cap L$ is observable in G, and our proof of Theorem 11 is complete.

There are some immediate algebraic-geometric consequences of our results that are worthy of note. *Let G be an irreducible algebraic linar group over an algebraically closed field, and let K and L be algebraic subgroups of G such that G/K and G/L are quasi-affine. Then $G/(K \cap L)$ is quasi-affine. Also, if K and L are algebraic subgroups of G such that $K \subset L$, and G/L and L/K are quasi-affine, then G/K is quasi-affine.* Indeed, in the first case, K and L are observable in G, by Theorem 4. Hence, by Theorem 11, $K \cap L$ is observable in G. By Theorem 4, $G/(K \cap L)$ is therefore quasi-affine. In the second case, we conclude from Theorem 4 that K is observable in L and that L is observable in G. This implies quite evidently that K is observable in G, so that, by Theorem 4, G/K is quasi-affine.

6. Some illustrative examples. There is a simple example, pointed out by Rosenlicht [3, p. 223], showing that G/K may be quasi-affine without being affine. Let F be an algebraically closed field, $G = SL(2, F)$, K the subgroup of all matrices of the form $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$. Then the algebraic variety G/K is isomorphic with the variety resulting from $F \times F$ when the point $(0, 0)$ is removed.

Let $G = GL(2, F)$, and let K be the algebraic subgroup consisting of all matrices of the form $\begin{pmatrix} a & b \\ 0 & a^2 \end{pmatrix}$, $a \neq 0$. It is not difficult to verify directly

that the condition of Theorem 1 is satisfied by (G,K), so that K is observable in G. On the other hand, (G,K) satisfies none of the intrinsic group-theoretical sufficient conditions of Theorem 2. Let L be the algebraic subgroup of G consisting of all matrices of the form $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$, $a \neq 0$. Then L is also observable in G. Let U be the group generated by K and L. Then U is actually the direct product $K \times L$ and consists of all non-singular triangular matrices. The algebraic variety G/U is complete (it is the well-known flag manifold), and thus is certainly not quasi-affine. Hence U is not observable in G, although it is the direct product of two observable subgroups.

More generally, let $G = GL(n,F)$, with F algebraically closed, and let K be the unipotent algebraic subgroup of all triangular matrices with 1's on the diagonal. Let L be the algebraic subgroup of G consisting of all diagonal matrices. Then K and L are both observable in G. On the other hand, the group U generated by K and L is a semidirect product $K \cdot L$, and consists of all non-singular triangular matrices. Since G/U is complete, U is not observable in G.

University of California, Berkeley
and Yale University.

REFERENCES.

[1] C. Chevalley, "Classification des groupes de Lie algébriques," *Séminaire Ecole Normale Supérieure*, 1956-1958, Exposé 8.

[2] ———, *Théorie des groupes de Lie*, Tome II, Paris, 1951.

[3] ———, ———, Tome III, Paris, 1955.

[4] G. Hochschild, "Cohomology of algebraic linear groups," *Illinois Journal of Mathematics*, vol. 5 (1961), pp. 492-519.

[5] G. Hochschild and G. D. Mostow, "Extension of representations of Lie groups and Lie algebras, I," *American Journal of Mathematics*, vol. 79 (1957), pp. 924-942.

[6] ———, "Representations and representative functions of Lie groups," *Annals of Mathematics*, vol. 66 (1957), pp. 495-542.

[7] G. D. Mostow, "Extensions of representations of Lie groups, II," *American Journal of Mathematics*, vol. 80 (1958), pp. 331-347.

[8] ———, "Fully reducible subgroups of algebraic groups," *American Journal of Mathematics*, vol. 78 (1956), pp. 200-221.

[9] M. Rosenlicht, "On quotient varieties and the affine embedding of certain homogeneous spaces," *Transactions of the American Mathematical Society*, vol. 101, pp. 211-223.

[10] M. Rosenlicht, "Some rationality questions on algebraic groups," *Annali di Matematica Pura ed Applicata*, vol. 43 (1957), pp. 25-50.

ON RADICALS OF DISCRETE SUBGROUPS OF LIE GROUPS.*

By Louis Auslander.[1]

1. Introduction. Zassenhaus [1] showed that subgroups H of the general linear group $GL(n, R)$ possess unique maximal normal solvable subgroups. (For a recent account of this result see [2].) He called the maximal normal solvable subgroup of H the radical of H. For our purpose it will be more convenient to call these subgroups discrete radicals. We will denote the discrete radical of H by $r(H)$. We will call a connected Lie group an analytic group. Let G be an analytic group and let H be a subgroup of G. Then since G has a matrix representation with the kernel of the representing homomorphism in the center of G, it follows that any subgroup H of G has a discrete radical, $r(H)$.

Now, in addition to the discrete radical of a group H, one may consider the more refined concept of the discrete nil-radical of H. By this we will mean the maximal normal nilpotent subgroup of H. From standard considerations of algebraic groups, it is easy to show that any subgroup H of an analytic group G has a discrete nil-radical, denoted by $n-r(H)$. Zassenhaus further stated a theorem which gives conditions under which $r(H), H \subset GL(n, R)$, is non-trivial. Unfortunately, the proof Zassenhaus gave his assertion is in error. (Zassenhaus has assured the author that his method of proof can be corrected to give a valid proof of his theorem.) The Zassenhaus result may be stated as follows:

Zassenhaus Lemma. *Let G be an analytic group and let A be a closed, normal, simply connected abelian analytic sumbroup of G. Let Γ be a discrete subgroup of G. If the image of in G/A is not discrete, then $r(\Gamma)$ is non-trivial. To be more precise, Γ's intersection with the identity component of the closure of the group ΓA is in $r(\Gamma)$.*

It will be our purpose in this paper to give a new proof of the Zassen-

* Received February 15, 1962.

[1] This research was supported by The National Science Foundation under Grants NSF 15565 and 18995 and the Office of Ordnance Research under contract SAR-DA 19-020-ORD-5254.

haus lemma and outline its central position in unifying certain studies concerning subgroups of Lie groups with compact quotient.

2. Zassenhaus lemma. We will now state the lemma in a slightly more convenient, but equivalent form to that in the introduction. For notational convenience we will use $I(\ )$ to denote the identity component of the closure of the group in the bracket.

ZASSENHAUS LEMMA. *Let G be a Lie group and let A be a closed, normal vector subgroup of G. Further let Γ be a discrete subgroup of G. Then $I(\Gamma A)$ is solvable and hence $\Gamma \cap I(\Gamma A) \subset r(\Gamma)$.*

Proof. If $I(\Gamma A) = A$, then the lemma is trivially true. Let K be $I(\Gamma A)/A$, dim. $K > 0$. For any elements a and b in a group, we will let (a, b) denote the commutator $aba^{-1}b^{-1}$. Let U_1 be an open set in K such that, for $k_1, k_2 \in U_1$, $(k_1, k_2) \in U_1$. If we choose a basis in A and let $ad(k) = \eta A \eta^{-1}$, where η is any pre-image of k, we may also require that for every $k \in U_1$ every entry in the matrix $(ad(k) - I)$ shall have absolute value less than 10^{-1}, where I is the identity matrix. We will denote the fact that every entry in the matrix $(ad(k) - \mathrm{I})$ has absolute value less than 10^{-r} by $| ad(k) - I | < 10^{-r}$. Let $U \subset U_1$ be such that if $k_1, k_2 \in U$ and $| ad(k_1) - I | < 10^{-r}$, $| ad(k_2) - I | < 10^{-s}$ then $| ad((k_1, k_2)) - I | < 10^{-t}$, where $t = \max.(r, s) + 1$ and $r > 0$, $s > 0$. The existence of the neighborhood U follows from Lemma 3.1 in [6].

For any $k \in (I(\Gamma A) \cap \Gamma A/A) \cap U$, let $\gamma \in \Gamma$ be a pre-image of k. Then $\gamma = xy$, where $y \in A$ and $x \in U^*$, where U^* is a neighborhood of the identity in G whose closure is compact and which maps onto U by the natural projection. We may require that U^* have the further property that if $u_1, u_2 \in U^*$ $(u_1, u_2) \in U^*$. Since Γ is a discrete subgroup of H, there exists a $d_0 > 0$ such that $| y | > d_0$ relative to the euclidean norm determined by the fixed basis already chosen in A. Then

$$(\gamma_1, \gamma_2) = (x_1, x_2) ad(x_1^{-1} x_2^{-1}) [(ad(x_2) - I) y_1 + (I - ad(x_1)) y_2]$$

where $\gamma_i = x_i y_i$, $i = 1, 2$, $ad(x_i) = ad(k_i)$, $k_i \in (I(\Gamma A) \cap \Gamma A/A) \cap U$. Now since $I(\Gamma A)/A$ is an analytic group, it is generated by any neighborhood of the identity and hence by the closure of the group generated by any dense subset of elements of Γ such that $\gamma_i = x_i y_i$, $x_i \in U^*$, $y_i \in A$, $i = 1, \cdots, r$. We will show that the group generated by the set $\gamma_1, \cdots, \gamma_r$ is nilpotent. This amounts to showing that any collection of $k > 0$ elements from $\gamma_1, \cdots, \gamma_r$

combined by commutators in any order is zero. But this follows easily from the previous discussion. For, if max. $|y_i| = D$, choose k such that $10^{-k}D < d_0/2$.

This proves that the group generated by any finite collection of elements in U^* is nilpotent. Hence, a group with the same Lie algebra as $I(\Gamma A)/A$ in some matrix representation is solvable. We can then conclude that $I(\Gamma A)$ is solvable.

PROPOSITION 2. *Let G be a Lie group, contain R as a closed normal solvable subgroup, and let R be connected and simply connected. Further let Γ be a discrete subgroup of G. Then $I(\Gamma R)$ is solvable.*

Proof. We will prove the theorem by induction on the number of steps of solvability of R. If R is abelian, we have proven the proposition in the Zassenhaus lemma. Assume the assertion is false. The there exist arbitrarily close to the identity in $I(\Gamma R)/R$ finite sets of elements in $I(\Gamma R) \cap \Gamma R/R$ which do not generate nilpotent groups. Now arguing as in the Zassenhaus lemma, we must have that the image of Γ in $G/[R,R]$ is not discrete. Let

$$R_2 = [R,R], \cdots, R_k = [R_{k-1}, R_{k-1}] \text{ and } R_{k+1} = e,$$

where e is the identity element of G. Then it is easy to see that the image of Γ in G/R_s is not discrete for any s and hence the image of Γ in G is not discrete. This contradicts the definition of Γ and proves our assertion.

3. Discrete radicals. We will adopt the following conventions: Let G be a simply connected analytic group and let the Levi decomposition of $G = R \cdot S$, where R is the radical of G (i.e. the maximal analytic normal solvable subgroup); S is a semi-simple analytic group; and, the dot denotes the semi-direct product. Now S is the direct product of simple analytic groups S_i, $i = 1, \cdots, k$. We will call the S_i the simple components of S. We will further use C to denote the direct product of all the compact components of S and call C the compact part of S. Clearly the group $R \cdot C$ is unique and independent of the semi-direct product used in its definition. In addition to the radical of G, we will need the more refined notion of the nil-radical of G. By the nil-radical, we mean the maximal normal nilpotent analytic subgroup of G. From standard considerations, it is easly to see that any subgroup H of an analytic group G has a maximal nilpotent normal subgroup called the discrete nil-radical and denoted by $n - r(H)$.

PROPOSITION 3. *Let G be a simply connected analytic group with Levi decomposition $G = R \cdot S$; let Γ be a discrete subgroup of G with compact quotient; and let C be the compact part of S. Then $I(\Gamma R)/R$ is abelian and contained in C.*

This Proposition follows easily from the Zassenhaus lemma and work of Borel [7].

THEOREM 1. *Let G be a simply connected analytic group with Levi decomposition $G = R \cdot S$ and let C be the compact part of S. If Γ is a discrete subgroup of G such that G/Γ is compact, then we have the following:*

1. *$\Gamma \cap R \cdot C$ has compact quotient in $R \cdot C$.*

2. *$\Gamma/\Gamma \cap R \cdot C$ is a discrete subgroup with compact quotient in S/C.*

3. *$r(G)C/r(G)C \cap \Gamma$ has a finite number of components and its identity component is compact.*

Proof. The proof of 1 and 2 of this theorem is equivalent to the proof that $I(\Gamma RC) = RC$. But this follows easily from Proposition 3 and the fact that C is compact. Hence all that remains is to prove 3.

Note first that the identity component of $r(G) = R$ and hence by 2 above if $D = (r(G)/RC) \cap (S/C)$ our assertion amounts to verifying that $(\Gamma/\Gamma \cap RC) \cap D$ is of finite index in D. Hence we lose no generality in considering just the following case: Let S be a connected, simply connected, semi-simple Lie group with no compact component and let Γ be a discrete subgroup of S such that S/Γ is compact. Let D be the discrete radical of S and let $C \subset D$ be the center of S. We first note that C is of finite index in D. For let S/C^*, $C^* \subset C$ have a faithful matrix representation. Then D/C^* is a normal solvable subgroup of a semi-simple matrix group. Then arguing as in Proposition 3, we see that L/C^* is finite and hence D/C is finite. We will now see that $\Gamma \cap C$ is of finite index. Assume this statement is false. Then C cannot be a finite group and further the group ΓC cannot be discrete as both S/Γ and $S/\Gamma C$ are compact and one is an infinite covering of the other. Then $I(\Gamma C)$ is non-trivial and there exist $\gamma_1, \gamma_2 \in \Gamma$ such that $\gamma_i = x_i c_i$, $i = 1, 2$, where x_i are arbitrarily close to the identity in S and $c_i \in C$. Hence $(\gamma_1, \gamma_2) = (x_1, x_2)$ is arbitrarily close to the identity and since Γ is discrete it must be the identity. Hence $I(\Gamma C)$ is an abelian analytic group. Hence in S/C^* the image I^* of $I(\Gamma C)$ must be an analytic abelian group with the property that $\gamma^* I^* \gamma^{*-1} = I^*$ for all $\gamma^* \in \Gamma C^*/C^*$. But this is impossible.

Hence $\Gamma \cap C$ is of finite index in C and since C is of finite index in D, $\Gamma \cap D$ must be of finite index in D. This proves our assertion.

COROLLARY. *Let G be a simply connected analytic group and let Γ be a discrete subgroup of G with compact quotient. Then G/Γ is a fiber bundle over a space $\mathfrak{X}$ with fiber Y, where*

$$Y = R \cdot C/\Gamma \cap R \cdot C \text{ and } \mathfrak{X} = (S/C)/(\Gamma/\Gamma \cap R \cdot C).$$

Remark. It would be worthwhile determining whether this bundle can ever be non-trivial and, if so, the structure of the group G in these cases.

4. Further applications. Zassenhaus related his result to the classical theorem of L. Bieberbach [3] on space groups, which in modern language may be restated as follows:

BIEBERBACH THEOREM. *Let Γ be a discrete subgroup of the group of rigid motions $R(n)$ of Euclidean n-dimensional space such that $R(n)/\Gamma$ is compact. Let T denote the subgroup of $R(n)$ consisting of pure translations. Then the identity component of the closure of the group ΓT is T.*

Now the already mentioned Bieberbach theorem is a refinement of Theorem 1 in a special case; i.e. it gives information about discrete subgroups of $R \cdot C$ with compact quotient. Another such refinement is the following weak form of a theorem of Mostow [4].

MOSTOW THEOREM. *Let Γ be a discrete subgroup with compact quotient of the connected, simply connected solvable analytic group R. Let N be the nil-radical of R. Then $I(\Gamma N) = N$ and $N/n - r(\Gamma) \cap N$ is compact.*

One could use the Zassenhaus lemma to give a new proof of the Mostow theorem. Let us state the following generalizations of the Mostow and Bieberbach theorems:

THEOREM 2. *Let G be a simply connected analytic group with no compact normal subgroup and let Γ be a discrete subgroup of G with compact quotient. Further let N be the nilradical of G. Then $I(\Gamma N) = N$ and $N/n - r(\Gamma) \cap N$ is compact.*

Theorem 2 follows easily from Theorem 1, the weak Mostow theorem, and the work in [5] and we will omit its proof. It should be noted that

Theorem 2 is false with the assumption of no compact normal subgroups removed.

YALE UNIVERSITY AND INDIANA UNIVERSITY.

REFERENCES.

[1] H. Zassenhaus, "Beweis eines Satze über diskrete Gruppen," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 19 (1938), pp. 289-312.

[2] I. Kaplansky, *Introduction to Differential Algebra*, Paris (1957).

[3] L. Bieberbach, "Über die Bewegungsgruppen der Euklidischen Raume I," *Mathematische Annalen*, vol. 70 (1911), pp. 297-336.

[4] G. D. Mostow, "Factor spaces of solvable groups," *Annals of Mathematics*, vol. 60 (1954), pp. 1-27.

[5] L. Auslander, "Bieberbach's theorem on space groups and discrete uniform subgroups of Lie groups, II," *American Journal of Mathematics*, vol. 83 (1961), pp. 276-280.

[6] ———, "Bieberbach's theorem on space groups and discrete uniform subgroups of Lie groups, I," *Annals of Mathematics*, vol. 71 (1960), pp. 579-590.

[7] A. Borel, "Density properties for certain subgroups of semi-simple groups without compact components," *Annals of Mathematics*, vol. 72 (1960), pp. 179-188.

ON THE SPECIALIZATION OF BIRATIONALLY EQUIVALENT CURVES.*

By IRWIN FISCHER.

It has been proved by Chow and Lang [1] that the birational equivalence of algebraic curves is preserved under specialization, provided that the specialized curves are non-singular. We shall prove this result also for singular curves, provided that the genus of each of the specialized curves equals that of the original curves. See van der Waerden [5] for the case of plane curves with ordinary double points.

By a curve we shall mean a positive 1-dimensional cycle in projective space, without multiple components. By the genus $g(Z)$ of an absolutely irreducible curve Z we shall mean the genus of its function field over any perfect field of definition. It does not depend on the particular perfect field chosen. In the following k will denote an arbitrary ground field. If Z is any curve $k(Z)$ will be the field generated over k by the ratios of the coefficients of the associated form of Z.

Later on we shall work with reducible curves. We refer to Rosenlicht [4] for the concepts of place, function ring, general point, etc. and for the Riemann-Roch Theorem for reducible curves.

We shall call a formal sum of points on a curve a 0 dim. cycle. The term divisor will always be used for sums of places. By the center of a divisor we mean the cycle which consists of the centers of the places in the divisor, counted with their multiplicities.

We shall make use of the following theorem, due to Hironaka. [3, p. 194, Th. 5.]

THEOREM 1. *Let the absolutely irreducible curve Z specialize over the field k to the curve $\bar{Z}=\bar{Z}_1+\bar{Z}_2+\cdots+\bar{Z}_r$, where $\bar{Z}_1, \bar{Z}_2, \cdots, \bar{Z}_r$ are the absolutely irreducible components of $\bar{Z}$. Then*

$$g(\bar{Z}_1)+g(\bar{Z}_2)+\cdots+g(\bar{Z}_r)\leqq g(\bar{Z}).$$

* Received June 30, 1961.

We now state our main result.

THEOREM 2. *Let Z_1, Z_2 be absolutely irreducible curves which are birationally equivalent to each other (over some common field of definition). Let (Z_1, Z_2) specialize over the field k to $(\bar{Z}_1, \bar{Z}_2)$, where $\bar{Z}_1$, $\bar{Z}_2$ are also absolutely irreducible, and are both of the same genus as Z_1, Z_2 Then $\bar{Z}_1$, $\bar{Z}_2$ are birationally equivalent.*

Proof. Let G be a birational correspondence between Z_1 and Z_2. Then G is itself an algebraic curve which is birationally equivalent to both Z_1 and Z_2. Therefore G is an absolutely irreducible curve of the same genus g as Z_1 and Z_2.

Extend the specialization given in the theorem to a specialization of G, yielding $(Z_1, Z_2, G) \to (\bar{Z}_1, \bar{Z}_2, \bar{G})$, the arrow denoting specialization over the field k.

Since specialization preserves the operation of algebraic projection, from $\mathrm{pr}_1 G = Z_1$ and $\mathrm{pr}_2 G = Z_2$ we deduce $\mathrm{pr}_1 \bar{G} = \bar{Z}_1$, $\mathrm{pr}_2 \bar{G} = \bar{Z}_2$. Hence either $\bar{G}$ is absolutely irreducible, in which case it is a birational correspondence between $\bar{Z}_1$ and $\bar{Z}_2$, or it has the form $\bar{P} \times \bar{Z}_2 + \bar{Z}_1 \times \bar{Q}$, where $\bar{P}$ is a point of $\bar{Z}_1$ and $\bar{Q}$ is a point of $\bar{Z}_2$. In this latter case, $\bar{G}$ has two components, each being a curve of genus g, neither of them being multiple components. By Theorem 1 we obtain $2g = g + g \leqq g$. If $g > 0$ this is impossible. This proves Theorem 2 for $g > 0$. For $g = 0$ the conclusion of the theorem is trivial.

Our result is an immediate consequence of Hironaka's inequality (Theorem 1). It may be of interest to see that, if $g > 1$, it follows from a weaker inequality, which can be proved in a simpler fashion.

THEOREM 3. *Let the absolutely irreducible curve Z specialize over the field k to the curve $\bar{Z} = \bar{Z}_1 + \bar{Z}_2 + \cdots + \bar{Z}_r$, where $\bar{Z}_1, \bar{Z}_2, \cdots, \bar{Z}_r$ are the absolutely irreducible components of $\bar{Z}$. Then*

$$g(\bar{Z}_1) + g(\bar{Z}_2) + \cdots + g(\bar{Z}_r) - r + 1 \leqq g(Z).$$

Proof. We can assume that k is an infinite field, by adjoining a transcendental if it is finite. Let Z, $\bar{Z}$ lie in a projective space of dimension n. We can then project Z, $\bar{Z}$ from a suitable $(n-3)$-dimensional subspace into a plane in such a way that the projection is birational on Z and the components of $\bar{Z}$. We can therefore assume without any loss of generality that Z, $\bar{Z}$ are plane curves.

We can give the set of all forms of degree h in the indeterminates X_0, X_1, X_2, with coefficients in the universal domain, the structure of a projective space V, by taking these coefficients to be projective coordinates.

Let U be the subspace of V which consists of all forms which vanish on Z. Let W be a complementary subspace to U in V, chosen in such a way that it is defined by homogeneous linear equations with coefficients in the prime field of our universal domain. Then if $m = \text{order}\, Z = \text{order}\, \bar{Z}$, we have for h large enough,

$$\begin{aligned}\dim W = h(h+3)/2 - (h-m+1)(h-m+2)/2 \\ = mh - (m+1)(m+2)/2.\end{aligned}$$

W is a maximal linear subspace of V which contains no form that vanishes on $\bar{Z}$. No form in W vanishes on Z. We can see this as follows. Let F be a form in W which vanishes on Z. Extend the specialization $Z \to \bar{Z}$ to one of F, obtaining $(Z, F) \to (\bar{Z}, \bar{F})$, over the field k. Since W is defined by linear homogeneous equations whose coefficients lie in k, we have $\bar{F} \in \bar{W}$. But $\bar{F}$ vanishes on $\bar{Z}$, which leads to a contradiction. Therefore no form in W vanishes on Z, and W is a maximal linear subspace with this property.

Let $m_i = \text{order}\, \bar{Z}_i$, $i = 1, 2, \cdots, r$. Let W_i be the linear subspace of W which consists of all forms in W which vanish on $\bar{Z}_i$. A simple calculation shows

$$\dim W_i = (m - m_i)h + m_i(m_i - 3)/2 - (m-1)(m-2)/2, \; i = 1, 2, \cdots, r.$$

The adjoint forms of order h in W cut out complete linear series on Z. (Gorenstein [2]). If we denote the degree of the adjoint divisor by v, it follows from the Riemann-Roch theorem that for h sufficiently high we have $mh - v - g + 1$ adjoint forms in W which are linearly independent over the field $k(Z)$. These forms determine a linear subspace S of W, with $\dim S = mh - v - g$.

Let C be the center on Z of the adjoint divisor of Z.

Extend the specialization $Z \to \bar{Z}$, over the field k, to $(Z, S, C) \to (\bar{Z}, \bar{S}, \bar{C})$.

From $\dim \bar{S} = mh - v - g$, we find that for sufficiently high h, $\bar{S}$ is not contained in W_i, $i = 1, 2, \cdots, r$. Hence it follows that almost all forms in $\bar{S}$ do not vanish on any component of $\bar{Z}$.

There are a finite number of divisors $\bar{A}_1, \bar{A}_2, \cdots, \bar{A}_q$ of the reducible curve $\bar{Z}$ whose centers on $\bar{Z}$ equal $\bar{C}$. Consider the set of all forms $\bar{g}$ in W such that $v_p(\bar{g}) \geqq v_p(\bar{A}_j)$ for all places p of $\bar{Z}$. These conditions being

linear we obtain for each j a linear subspace W_j^* of W, $j = 1, 2, \cdots, q$. Those members of $\bar{S}$ which intersect $\bar{Z}$ properly belong to $\bigcup_{j=1}^{q} W_j^*$. Since almost all members of $\bar{S}$ (namely those not in $\bigcup_{i=1}^{r} W_i$) intersect $\bar{Z}$ properly, we have $\bar{S} \subset \bigcup_{j=1}^{q} W_j^*$. Therefore $\bar{S} \subset W_j^*$ for some j. Dropping the subscript, we call the corresponding divisor $\bar{A}$, and see that $v_p(\bar{g}) \geqq v_p(\bar{A})$ for every form $\bar{g}$ in $\bar{S}$. We may, of course, have $v_p(\bar{g}) = \infty$.

Let $\bar{k}$ be the algebraic closure of $k(\bar{Z})$. Let

$$\bar{g}_0, \bar{g}_1, \cdots, \bar{g}_s \qquad (s = mh - v + g)$$

be a linearly independent set of forms in $\bar{S}$ such that $\bar{g}_0$ intersects $\bar{Z}$ properly. Let $\bar{E}$ be the intersection divisor of the form $\bar{g}_0$ with $\bar{Z}$. Then if (y_0, y_1, y_2) is a homogeneous general point of $\bar{Z}$ over $\bar{k}$, set

$$\phi_j = \bar{g}_j(y_0, y_1, y_2)/\bar{g}_0(y_0, y_1, y_2), \qquad j = 0, 1, \cdots, s.$$

Let $\bar{D} = \bar{E} - \bar{A}$. We have proved above that $v_p(\bar{g}_0) \geqq v_p(\bar{A})$ for every place p of $\bar{Z}$. Hence $\bar{D}$ is an integral divisor. Also $v_p(\bar{g}_j) \geqq v_p(\bar{A})$, for every p and j.

Let $L(\bar{D})$ be the vector space of all functions ϕ in the function ring of $\bar{Z}$ satisfying $v_p(\phi) \geqq -v_p(\bar{D})$ for all places p of $\bar{Z}$. Then $\phi_0 = 1, \phi_1, \phi_2, \cdots, \phi_s \in L(\bar{D})$, hence $\dim L(\bar{D}) \geqq mh - v - g + 1$. By the Riemann-Roch theorem for reducible curves, we have $\dim L(\bar{D}) = d(\bar{D}) - \bar{g} + 1 + i(\bar{D})$, with $\bar{g} = g(\bar{Z}_1) + g(\bar{Z}_2) + \cdots + g(\bar{Z}_r) - r + 1$.

For h sufficiently high, $\bar{D}$ is a non-special divisor since the degree of its projection on $\bar{Z}_j$ can then be made larger than $2g(\bar{Z}_j) - 2$, for each j. Hence if this is done we have $i(\bar{D}) = 0$, and it then follows that $mh - v - \bar{g} + 1 \geqq mh - v - g + 1$, or $\bar{g} \leqq g$. This is the desired conclusion.

If $g > 1$, Theorem 2 follows from Theorem 3 by the same argument used previously to deduce it from Theorem 1.

We have stated our result on the specialization of birationally equivalent curves only for specialization over a field. Hironaka proves Theorem 1 for specializations over a discrete, rank 1 valuation ring of finite type. Our proof of Theorem 3 holds for specializations over any discrete, rank 1 valuation ring. From these remarks it follows that if $g > 1$, Theorem 2 holds for specializations over discrete, rank 1 valuation rings, while if $g = 1$ it holds at least under the additional assumption that the rings are of finite type.

For non-singular curves, Chow and Lang have proved the theorem for specializations over any discrete, rank 1 valuation ring.

UNIVERSITY OF COLORADO.

REFERENCES.

[1] W. Chow and S. Lang, "On the birational equivalence of curves under specialization," *American Journal of Mathematics*, vol. 79 (1957), pp. 649-652.

[2] D. Gorenstein, "An arithmetic theory of adjoint plane curves," *Transactions of the American Mathematical Society*, vol. 72 (1952), pp. 414-436.

[3] H. Hironaka, "On the arithmetic genera and the effective genera of algebraic curves," *Memoirs of the College of Science, University of Kyoto, Series A*, vol. 30, No. 2 (1957), pp. 177-195.

[4] M. Rosenlicht, "Equivalence relations on algebraic curves," *Annals of Mathematics*, vol. 56 (1952), pp. 169-191.

[5] B. L. van der Waerden, "Zur algebraischen geometrie, XI, Projective und birationale aquivalenz und moduln von ebenen Kurven," *Mathematische Annalen*, vol. 114 (1937), pp. 683-699.

COHOMOLOGY OPERATIONS WITH LOCAL COEFFICIENTS.* [1]

By SAMUEL GITLER.

0. Introduction. The study of cohomology operations has been one of the important areas of research in Algebraic Topology for the last twelve years. They have been applied extensively to compute obstructions [18], to study the homotopy type of complexes [22] and to show essentiality of maps of spheres into spheres [3].

One of the basic operations is the reduced powers of Steenrod $\mathcal{P}^i$ [1], [19], [20]. If X is any space, G a cyclic group of prime order p, then $\mathcal{P}^i$ is a natural homomorphism.

$$\mathcal{P}^i\colon H^q(X;G) \to H^{q+2i(p-1)}(X;G)$$

$i = 0, 1, 2, \cdots$, where $H^r(X;G)$ is the r-th dimensional cohomology group of X with coefficients in G.

Steenrod [16], [17] and Reidemeister [15] independently generalized the concept of cohomology with coefficients in a group G to cohomology with coefficients in a local system of groups with basic group G; the local system of groups is characterized by an action of the fundamental group as automorphisms of G.

In this paper, the definition and theory of the reduced power operations $\mathcal{P}^i$ is extended to the case of cohomoloy with coefficients in a local system of groups whose basic group is a cyclic group of prime order $p \neq 2$.

All of the standard properties, when suitably interpreted hold in the extended theory. Furthermore there is only one extension which preserves these properties.

Examples are given of spaces which are distinguished by the $\mathcal{P}^i$ with non-simple coefficients, but which other known cohomology operations fail to distinguish.

Eilenberg [6] has shown that the cohomology theory with a local system of groups is naturally eqiuvalent to a theory of equivariant cohomology in the universal covering space with ordinary coefficients.

* Received April 9, 1962.

[1] This paper contains the main results of a Doctoral Thesis developed by the author under the guidance of Prof. N. E. Steenrod and submitted to Princeton University in 1960.

The construction of the $\mathfrak{P}^i$ is done with respect to this equivariant cohomology, and is based on the p-fold cartesian product of a space X, and the cyclic group of permutations of the factors, just as in the case of simple coefficients.

A second approach to cohomology operations with simple coefficients is provided by the Eilenberg-MacLane complexes $K(G, q)$ [4], [7]. A well known theorem establishes a one to one correspondence between cohomology operations and cohomology classes of $K(G, q)$.

The complexes $K(G, q)$ are readily generalized to study cohomology operations with local systems of groups by this method. The generalized Eilenberg-MacLane complexes $L_\pi(G, q)$ are complexes with only two non-vanishing homotopy groups, one of them the fundamental group which is π, the other is G in dimension q. The k-invariant is zero, but the complex is not a product because the fundamental group of the base operates non-trivially in the fibre.

Finally we show that the set of cohomology operations with initial and terminal local systems of groups with basic group cyclic of order $p > 2$, has a basis of: addition, cup products, Bockstein and $\mathfrak{P}^i$, which is the exact analogue of the results with ordinary coefficients (cf. [4]).

This paper has been divided into four chapters. Chapter I reviews briefly the cohomology theory with local coefficients, and its relation to an equivariant cohomology of the universal covering space. In chapter II we define the $\mathfrak{P}^i$ and state their basic properties. Chapter III deals with the generalized Eilenberg-MacLane complexes $L_\pi(G, q)$ and cohomology operations in general. Chapter IV contains the proofs of some of the theorems of chapters II and III.

Chapter I.

Cohomology with local coefficients and covering spaces.

1. Cohomology with local coefficients. By a space X, we will understand a space which is semilocally one connected,[2] pathwise connected and locally pathwise connected.

We assume we are given a local system of groups $\mathcal{G} = \{G_x\}$ on X as defined by Steenrod [16]. A mapping $f\colon Y \to X$ induces a local system of groups $f^*\mathcal{G}$ on Y. We may then consider the cohomology groups $H^*(X; \mathcal{G})$ and $H^*(Y; f^*\mathcal{G})$ and the induced mapping:

[2] See [16], p. 64.

$$f^*: H^*(X;\mathcal{G}) \to H^*(Y;f^*\mathcal{G}).$$

In this paper we consider always simplicial singular cohomology [6].

Often it is more convenient to restrict the local system of groups to a local group in the sense of Olum [13]. They are naturally equivalent.

We will consider both methods in this paper.

2. **Pairings of local systems and cup products.** Let $\mathcal{G}_i$ be a local system of groups over the space X_i $(i=1,2)$. Then in an obvious way $\mathcal{G}_1 \otimes \mathcal{G}_2 = \{(G_1 \otimes G_2)_{x_1,x_2}\}$ is a local system of coefficients over the cartesian product $X_1 \times X_2$. If $X_1 = X_2 = X$, denote by $\mathcal{G}_1 \circ \mathcal{G}_2$ the induced local system of groups on X corresponding to the diagonal mapping $d: X \to X \times X$. $\mathcal{G}_1 \circ \mathcal{G}_2$ is called the *product* of $\mathcal{G}_1$ and $\mathcal{G}_2$. $\mathcal{G}^n$ will denote the product of $\mathcal{G}$ with itself n-times.

One is interested in the isomorphism classes of local systems of groups. In this sense, the above product is associative and commutative. Moreover if $\mathcal{G}_1$, $\mathcal{G}_2$ are local systems of groups, then we have defined a cup product pairing [16]:

$$H^p(X;\mathcal{G}_1) \otimes H^q(X;\mathcal{G}_2) \to H^{p+q}(X;\mathcal{G}_1 \circ \mathcal{G}_2) \tag{2.1}$$

which is natural, associative and commutative (in the sense of graded objects).

We will be particularly interested in local systems whose basic group is a cyclic group of prime order $p \neq 2$.

2.2. Let $\mathfrak{Z}_p(X)$ denote the set of all isomorphism classes of local systems of groups over X with basic group a cyclic group of order p. By $H^*(X;\mathfrak{Z}_p)$ we will denote the bigraded group with bidegrees $(q,\mathcal{G})$ where q is an integer $\geqq 0$ and $\mathcal{G} \in \mathfrak{Z}_p(X)$.

Every class in $\mathfrak{Z}_p(X)$ has a representative whose basic group is the additive group of Z_p, the integers modulo p.

2. 3. *$\mathfrak{Z}_p(X)$ is an abelian group under the product $\mathcal{G}_1 \circ \mathcal{G}_2$. Its unit is (the class of) the trivial local system of groups, and every element $\mathcal{G} \in \mathfrak{Z}_p(X)$ has finite order m, where m divides $p-1$.*

From the definition of the product, it is clear that it passes to classes. Associativity and commutativity follow from the corresponding properties in the tensor product. Thus $\mathfrak{Z}_p(X)$ is an abelian monoid with unit the class of the ordinary coefficients Z_p. We may consider a representative $\mathcal{G}_0$ of a class in $\mathfrak{Z}_p(X)$ to have basic group Z_p. By Section 1, $\mathcal{G}_0$ is uniquely determined by a local group G_{x_0}, i.e. a homomorphism: $\pi_1(X,x_0) \to A(Z_p)$ where $A(Z_p)$

is the automorphism group of G_{x_0}. Now $A(Z_p)$ is cyclic of order $p-1$, it then follows that $\mathfrak{F}_p(X)$ is in 1-1 correspondence with homomorphisms

$$\pi_1(X) \to A(Z_p)$$

i.e. $\mathfrak{F}_p(X) \leftrightarrow \mathrm{Hom}(H_1(X), A(Z_p))$ and in such a way that this is an isomorphism of groups.

2.4. If $\{\mathfrak{G}\}$ denotes the subgroup generated by $\mathfrak{G}$ in $\mathfrak{F}_p(X)$, $H^*(X;\{\mathfrak{G}\})$ is a subgroup of $H^*(X;\mathfrak{F}_p)$ and we have:

$H^*(X;\mathfrak{F}_p)$ *is a commutative, associative, bigraded ring under cup products; for every* $\mathfrak{G} \in \mathfrak{F}_p(X), H^*(X;\{\mathfrak{G}\})$ *is a subring; and if* $f: Y \to X$ *is a mapping, then* f *induces a ring homomorphism*:

$$f^*: H^*(X;\mathfrak{F}_p) \to H^*(Y;\mathfrak{F}_p)$$

so that if $\mathfrak{G} \in \mathfrak{F}_p(X)$,

$$f^*: H^*(X;\mathfrak{G}) \to H^*(Y;f^*\mathfrak{G})$$

The above statements follow easily from the properties of cup products and (2.2).

3. Covering spaces and local systems of groups. In this paragraph we obtain some results relating the cohomology with local coefficients on a space X and the ordinary cohomology of its universal covering space.

We now consider spaces with base points, so that mappings preserve base point.

Let X be a space, x_0 its base point. Then $\tilde{X}$ will denote the universal covering space of X based at x_0. (cf. [6]). $\tilde{X}$ is a covariant functor in the category of spaces with base point. Let $p: \tilde{X} \to X$ denote the projection.

If now $\mathfrak{G}$ is a local system of groups on X, G_0 the local group at x_0 then $\pi_1(X, x_0)$ operates in $\tilde{X}$ as group of covering translations, and on G_0. Relative to this action we may form $H_e{}^q(\tilde{X}; G_0)$ the equivariant cohomology groups of $\tilde{X}$ with coefficients in G_0. For their definition and the following main theorem we refer to [6]:

3.1. Theorem (Eilenberg). *The projection* $p: \tilde{X} \to X$ *induces an isomorphism*:

$$p^*: H^q(X;\mathfrak{G}) \to H_e{}^q(\tilde{X}; G_0)$$

for all q, *which is natural with respect to mappings of spaces with base point.*

3.2. Now assume that $\pi = \pi_1(X, x_0)$ is finite of order m. We have a natural mapping:

$$(3.3) \qquad i : H_e^*(\tilde{X}; G_0) \rightarrow H^*(\tilde{X}; G_0)$$

induced by the inclusion of cochains and a transfer homomorphism (cf. [5], [21]).

$$(3.4) \qquad \mathrm{Tr}: H^*(\tilde{X}; G_0) \rightarrow H_e^*(\tilde{X}; G_0)$$

defined in the cochains by: $\mathrm{Tr}(f) = \sum_{s \in \pi} sf$ so that in cohomology:

$$(3.5) \qquad \mathrm{Tr}\ (x) = mx.$$

Now π operates in $H^*(\tilde{X}; G_0)$, so let:

$$H^*(\tilde{X}; G_0)^{\pi} = \{x \in H^*(\tilde{X}; G_0) \mid sx = x \text{ for all } s \in \pi\}.$$

Then the following proposition is true:

3.6. *If multiplication by m is an automorphism of G_0, then*

$$i : H_e^*(\tilde{X}; G_0) \rightarrow H^*(\tilde{X}; G_0)$$

is a monomorphism, and its image is $H^*(\tilde{X}; G_0)^{\pi}$.

Proof. Multiplication by m in the coefficient group G_0 induces an automorphism $\eta: H_e^*(\tilde{X}; G_0) \rightarrow H_e^*(\tilde{X}; G_0)$ with $\eta(x) = mx$ for all x. Hence $(\eta^{-1}\mathrm{Tr})\, i(x) = x$ implies i is a monomorphism.

Also if $x \in H_e^*(\tilde{X}; G_0)$, z a representative cocycle of x, then $i\,(x)$ is also represented by z; but since $sz = z$ for all s, we have $i\,(x) \in H^*(\tilde{X}; G_0)^{\pi}$.

Conversely if $y \in H^*(\tilde{X}; G_0)^{\pi}$ and w is a representative cocycle, then $sw \sim w$ for all $s \in \pi$, hence in cochains:

$$i\ \mathrm{Tr}(w) = \sum_{s \in \pi} sw \sim mw, \text{ therefore } i\ \mathrm{Tr}(y) = my$$

But by the assumption we may divide by m, and thus $y \in im\, i$. (3.1) and (3.6) thus give:

3.7. *Let $\mathcal{G}$ be a local system of groups on X, G_0 the local group at x_0, $\tilde{X}$ the universal covering space of X at x_0. Then if $\pi = \pi_1(X, x_0)$ is of finite order m so that multiplication by m is an automorphism of G_0 then $p: \tilde{X} \rightarrow X$ induces a natural isomorphism:*

$$p^*: H^*(X; \mathcal{G}) \rightarrow H^*(\tilde{X}; G_0)^{\pi}.$$

Chapter II.

Reduced powers in equivariant cohomology.

4. Definition of the reduced powers. In this section we extend the definition of the Steenrod reduced powers to equivariant cohomology. We recall that the definition of the reduced powers with ordinary coefficients in a space X (cf. [2], [21]) is based in a "chain approximation" to the diagonal $d\colon X \to X^n$ which behaves properly with respect to the action of a permutation group π of the factors of X^n, where X^n is the n-fold cartesian product of X with itself.

Let ρ be a group, then by a ρ-space X we will mean a space X on which ρ acts without fixed points.

We will start with a ρ-space X, then X^n is also a ρ-space, under the diagonal action:

$$s(x_1, \cdots, x_n) = (sx_1, \cdots, sx_n)$$

where $s \in \rho$, $(x_1, \cdots x_n) \in X^n$. If π is a group of permutations of the factors of X^n, the action of ρ and π commute, so that $\pi \times \rho$ is a group of operators in X^n. The construction of the reduced powers with equivariant cohomology will then be based in a chain approximation to $d\colon X \to X^n$ which behaves properly with respect to the action of $\pi \times \rho$ on X^n.

Special assumption. Throughout this section, the coefficient group G will always be a cyclic group of order p (p prime > 2).

Let X be a ρ-space, G a ρ-module, then we may consider as before the equivariant cohomology $H_e{}^q(X; G)$.

For every q, let $M(q)$ be the elementary cochain complex having only one non-zero cochain group in dimension q, with generator u_1 of order p.

Then to every class $u \in H_e{}^q(X; G)$ we associate a class $C(u)$ of cochain homotopic mappings: $M(q) \to C_e{}^*(X; G)$, so that if $f \in C(u)$, $f^*(u_1) = u$. The elements of $C(u)$ are called cochain representations of u.

Let π be a subgroup of the symmetric group of degree p, then the construction of the π-reduced powers of a class $u \in H_e{}^q(X; G)$ is based in the following diagram (compare with [21]):

$$(4.1) \qquad W \otimes_\pi M^p \xrightarrow{\psi} W \otimes_\pi (C_e{}^*(X; G))^p \xrightarrow{\phi} C_e{}^*(X; G)$$

where we need to explain the undefined terms:

W is a π-free acyclic complex, $M = M(q)$, M^p and $(C_e{}^*(X; G))^p$ are the p-fold tensor products of M and $C_e{}^*(X; G)$ respectively.

$W\otimes(C_e^*(X;G))^p$ is the tensor product of a chain complex by a cochain complex, and is a cochain complex as defined by Steenrod. M^p is a cochain complex with only one non-zero cochain group in dimension pq, which is cyclic of order p, generated by $u_1{}^p$. π operates in M^p by the sign of the permutation if q is odd, and trivially if q is even. The action of π in $W\otimes M^p$ is the diagonal action. In $(C_e^*(X:G))^p$ the operations of π are defined by:

$$\alpha(v_1\otimes\cdots\otimes v_p)=\epsilon(\alpha)v_1\otimes\cdots\otimes v_{i+1}\otimes v_i\otimes v_{i+2}\otimes\cdots\otimes v_p\ .$$

where $\epsilon(\alpha)=(-1)^{\dim v_i \dim v_{i+1}}$ if α is the simple interchange of the i-th and $(i+1)$-factor. We remark that

$$\alpha(u_1\otimes\cdots\otimes u_p)\cdot(\sigma)=u_1\otimes\cdots\otimes u_p\cdot(\alpha^{-1}\sigma).$$

In $W\otimes(C_e^*(X;G))^p$, π acts diagonally.

As to the mappings, ψ is induced by a cochain representation

$$f\colon M\to C_e^*(X;G),\ \text{i.e. } \psi=1\otimes f^p.$$

Now ϕ is to be dual to a chain mapping:

$$\phi'\colon W\otimes C_*(X)\to(C_*)X))^p$$

which is $\pi\times\rho$-equivariant, where $\pi\times\rho$ acts on $W\otimes C_*(X)$ so that π acts in W, ρ on $C_*(X)$, while the action of $\pi\times\rho$ on $(C_*(X))^p$ is the one induced by the action of $\pi\times\rho$ on X^p. The existence and uniqueness of ϕ' is provided by

4.2. *Let X be a ρ-space, π a permutation group of degree p. Then there exists a $\pi\times\rho$-equivariant chain mapping:*

$$\phi'\colon W\otimes C_*(X)\to(C_*(X))^p$$

satisfying:

$$\phi'(w_0\otimes x_0)=x_0\otimes\cdots\otimes x_0$$

where w_0, x_0 are vertices of W and X respectively.

If ϕ'' is another such mapping, then ϕ' is $\pi\times\rho$ equivariantly homotopic to ϕ''.

The proof of this and the following proposition is obtained by using the theory of acyclic models and will be given in Section 9 (9.1).

The mapping ϕ' also satisfies a naturality condition. Let (X,ρ) be a pari consisting of a group ρ and a ρ-space X. A mapping:

$$f\colon (X,\rho)\to(Y,\rho')$$

consists of a mapping: $X \to Y$ and a homomorphism $\rho \to \rho'$ both denoted by f so that:

$$f(sx) = f(s)f(x).$$

A mapping ϕ' such as in (4.2) is called a *diagonal approximation*:

4.3. *Let f: $(X,\rho) \to (Y,\rho')$ be a mapping, and let ϕ_1, ϕ_2 be diagonal approximations in X and Y respectively. Then f induces a diagram*

$$\begin{array}{ccc} W \otimes C_*(X) & \xrightarrow{\phi_1} & (C_*(X))^p \\ \Big\downarrow 1 \otimes f_\# & & \Big\downarrow (f_\#)^p \\ W \otimes C_*(Y) & \xrightarrow{\phi_2} & (C_*(Y))^p \end{array}$$

which is commutative up to a homotopy which is equivariant relative to the mapping $1 \times f$: $\pi \times \rho \to \pi \times \rho'$.

Let now ϕ': $W \otimes C_*(X) \to (C_*(X))^p$ be a diagonal approximation as in (4.2); define

$$\phi_0\colon W \otimes (C^*(X;G))^p \to C^*(X;G^p)$$

by:

$$(4.4) \quad \phi_0(w \otimes u_1 \otimes \cdots \otimes u_p) \cdot \sigma = (-1)^{i(i-1)/2} u_1 \otimes \cdots \otimes u_p \cdot \phi'(w \otimes \sigma)$$

where $i = \dim w$, $u_1 \otimes \cdots \otimes u_p$ is an n-cochain of $(C^*(X;G))^p$, σ is an $(n-i)$-simplex of X, G is a ρ-module and G^p is the p-fold tensor product of G with itself.

The equivariance of ϕ' gives in a straight-forward manner that ϕ_0 is actually defined in $W \otimes_\pi (C^*(X;G))^p$.

Now the inclusion:

$$i\colon C_e^*(X;G) \to C^*(X;G)$$

induces a natural mapping

$$j\colon W \otimes_\pi (C_e^*(X;G))^p \to W \otimes_\pi (C^*(X;G))^p$$

and again it is not hard to see that $\phi_0 j$ sends $W \otimes_\pi (C_e^*(X;G))^p$ into $C_e^*(X;G^p)$. Thus we obtain

$$\phi_0 j\colon W \otimes_\pi (C_e^*(X;G))^p \to C_e^*(X;G^p)$$

which is a cochain mapping as may be easily verified.

But now if G is cyclic of order p, the mapping

$$\text{(4.5)} \qquad \mathrm{X}\colon G^p \to G$$

defined by picking a generator T of G and setting

$$\mathrm{X}(T \times \cdots \times T) = T$$

is an isomorphism, independent of the choice of T. Furthermore X is natural with respect to isomorphisms of cyclic groups of order p, it induces then a natural isomorphism:

$$\mathrm{X}\colon C_e^*(X; G^p) \to C_e^*(X; G)$$

and $\phi = \mathrm{X}\phi_0 j$.

We now return to (4.1) having defined all the terms, and we define:

$$\text{(4.6)} \qquad \Phi\colon H^*(W \otimes_\pi M^p) \to H_e^q(X; G)$$

to be the composition $(\phi\psi)^*$.

The Φ-image of $H^*(W \otimes_\pi M^p)$ is called the set of *π-reduced powers of the class $\bar{u}$.*

There were several choices made at various stages, and now we will see the degree of freedom involved in these choices.

4.7. $\Phi\colon H^*(W \otimes_\pi M^p) \to H_e^*(H; G)$ *is independent of*

1) *The cochain representation of $\bar{u}$*
2) *The diagonal approximation ϕ'*
3) *The particular π-free acyclic complex W.*

The proof of 1) is that of Steenrod in [20].

As for 2), let ϕ'' be another diagonal approximation, then by (4.2) ϕ' and ϕ'' are $\pi \times \rho$ equivariantly homotopic, i.e. there exists a mapping:

$$D'\colon W \otimes C_*(X) \to (C_*(X))^p$$

of degree $+ 1$ with:

$$\partial D' + D'\partial = \phi' - \phi''$$

and D' is $\pi \times \rho$-equivariant. If we now define

$$D'_0\colon W \otimes (C^*(X; G))^p \to C^*(X; G^p)$$

by:

$$D'_0(w \otimes v_1 \otimes \cdots \otimes v_p) \cdot (\sigma) = (-1)^{i(i-1)/2} v_1 \otimes \cdots \otimes v_p \cdot D'(w \otimes \sigma)$$

then as in the case of ϕ', D'_0 passes to give:

$$D_0\colon W \otimes_\pi (C_e^*(X; G^p) \to C_e^*(X; G^p)$$

Then if we set $D = \mathrm{X}D_0$, one easily verifies that D is a cochain homotopy connecting ϕ and ϕ_0, duals of ϕ' and ϕ''. We finally prove 3) as follows:

Let π' be a subgroup of π, V a π'-free acyclic complex. It is well known that the inclusion $\pi' \to \pi$ is covered by a unique homotopy class of π'-equivariant mappings $\bar{g}\colon V \to W$. Let $\phi'\colon W \otimes C_*(X) \to (C_*(X))^p$ be a $\pi \times \rho$-equivariant diagonal approximation, then $\phi_1 = \phi'(\bar{g} \otimes 1)$ is a $\pi' \times \rho$-equivariant diagonal approximation and we obtain a commutative diagram:

(4.8)

$$\begin{tikzcd} V \otimes_{\pi'} (C_e^*(X;G))^p \arrow[d, "\bar{g}\otimes 1"'] \arrow[dr, "\phi_1"] & \\ W \otimes_\pi (C_e^*(X;G))^p \arrow[r, "\phi"'] & C_e^*(X;G) \end{tikzcd}$$

From (4.8) one clearly obtains 3) plus the important fact that (4.8) in cohomology tells us that all π-reduced powers are $S(p)$-reduced powers, where $S(p)$ is the symmetric group of degree p.

We study now the behaviour of reduced powers with respect to mappings. Let Φ_u denote the mapping (4.6) defining the π-reduced powers of $\bar{u}$.

4.9. *Let $f\colon (Y, \rho') \to (X, \rho)$ be a mapping, G a ρ-group $u \in H_e^q(X;G)$. Then f induces a commutative triangle:*

$$\begin{tikzcd} & H_e^*(X;G) \arrow[d, "f^*"] \\ H^*(W \otimes_\pi M^p) \arrow[ur, "\Phi_u"] \arrow[r, "\Phi_v"'] & H_e^*(Y;G) \end{tikzcd}$$

where $v = f^* u$.

Proof. In the cochains we have:

(4.10)

$$\begin{tikzcd} & W \otimes_\pi (C_e^*(X;G))^p \arrow[r] \arrow[d, "V \otimes_\pi (f^\#)^p"] & C_e^*(X;G) \arrow[d, "f^\#"] \\ W \otimes_\pi M^p \arrow[ur, "\psi_u"] \arrow[r, "\psi_v"'] & W \otimes_\pi (C_e^*(Y;G))^p \arrow[r] & C_e^*(Y;G) \end{tikzcd}$$

and the triangle can be chosen commutative by (4.7). It remains to study the right square. If one considers the dual diagram:

$$\begin{array}{ccc} W \otimes C_*(Y) & \longrightarrow & (C_*(Y))^p \\ \Big\downarrow 1 \otimes f_\# & & \Big\downarrow (f_\#)^p \\ W \otimes C_*(X) & \longrightarrow & (C_*(X))^p \end{array}$$

then (4.3) asserts it is commutative up to $1 \times f$-invariant homotopy. The dual of this homotopy makes the square in (4.10) commutative up to homotopy.

4.11. We remark that there is a natural isomorphism:

$$\gamma\colon H_k(\pi; Z_p^{(q)}) \to H^{pq-k}(W \otimes_\pi M^p)$$

where $Z_p^{(q)}$ denotes the integers modulo p, and π acts in $Z_p^{(q)}$ by the sign of the permutation if q is odd and trivially if q is even.

If we take π to be the cyclic group of order p, $p \neq 2$, then every permutation is even, so that the action on $Z_p^{(q)}$ is trivial. We may then consider the inclusion in homology:

$$h_k\colon H_k(\pi; Z_p) \to H_k(S(p); Z_p^{(q)})$$

then Steenrod [19] shows that if q is even $h_{2k} = h_{2k-1} = 0$ if $2k$ is not an even multiple of $p-1$, and if q is odd, $h_{2k} = h_{2k-1} = 0$ if $2k$ is not an odd multiple of $p-1$. An immediate consequence of this is Thom's theorem:

4.12. THEOREM. *If p is an odd prime, π = cyclic group of order p, then $\Phi_{\bar u}^*(H^{pq-k}(W \otimes_\pi M^p)) = 0$ unless the change in dimension $(pq-k)-q$ is congruent to 0 or $1 \bmod 2(p-1)$.*

Another reduction in the study of cyclic reduced powers with ordinary coefficients is the relation between even and odd reduced powers:

Let $\xi \in H_k(\pi; Z_p)$, $\bar u \in H_\theta^q(X; G)$, then if we set $\xi(\bar u) = \Phi_{\bar u}^*(\gamma(\xi))$ the relation in question is:

(4.13) $$(\beta_* \xi) = \beta^*(\xi(\bar u))$$

(cf. [19], p. 217).

Here β_* is the homology Bockstein and β^* is the cohomology Bockstein associated with the exact sequence:

(4.14) $$0 \to Z_p \to Z_{p^2} \to Z_p \to 0.$$

We now introduce Bocksteins in equivariant cohomology and show that (4.13) holds with respect to equivariant reduced powers. Let

$$0 \to G' \to G \to G'' \to 0$$

be an exact sequence of ρ-modules. If X is a ρ-space, then we have an exact triangle:

$$\begin{tikzcd}
H_e^*(X;G') \arrow[rr, "i"] & & H_e^*(X;G) \arrow[dl, "j"] \\
& H_e^*(X;G'') \arrow[ul, "\delta^*"] &
\end{tikzcd}$$

where i, j are of degree 0, δ^* of degree $+1$.

If now Z_p is a ρ-module, then (4.14) can be made into a ρ-sequence by lifting the action of ρ on Z_p to an action of ρ on Z_{p^2} as follows:

Let $A(Z_p)$, $A(Z_{p^2})$ denote the automorphism groups of Z_p and Z_{p^2} respectively. Let k be a primitive root of the prime p, and T the generator of $A(Z_p)$ corresponding to k, i.e., $T(1) = k$. Define

$$L: A(Z_p) \to A(Z_{p^2})$$

by

$$L(T)(1) = k^p \operatorname{mod} p^2.$$

Then L is a homomorphism so that $T_1 = L(T)$ acts in pZ_{p^2} compatibly with the injection $0 \to Z_p \to Z_{p^2}$.

The lifting L is independent of the choice of the primitive root k of p. For let k' be another root and T' the corresponding generator, then $T' = T^s$ for some s with $(s, p-1) = 1$, so $k' \equiv k^s \operatorname{mod} p$ i.e. $k' \equiv k^s + lp \operatorname{mod} p^2$, and then if L' is the lifting associated to k',

$$L'(T')(1) = k'^p \equiv (k^s + lp)^p \equiv k^{sp} \equiv L(T^s)(1) \operatorname{mod} p^2$$

With this lifting L we make Z_{p^2} into a ρ-module and obtain in the usual way an equivariant Bockstein β^*:

$$\beta^*: H_e^q(X; Z_p) \to H_e^{q+1}(X; Z_p) \tag{4.15}$$

for any ρ-space X, where the action of ρ is the *same* in the initial and terminal coefficient. It is easily verified that β^* is natural with respect to mappings $(X, \rho) \to (X', \rho')$.

4.16. *The cyclic reduced powers satisfy the relation*:

$$(\beta_*\xi)(\bar{u}) = \beta^*(\xi(\bar{u}))$$

for any class $\xi \in H_*(\pi; Z_p)$ *and any* $\bar{u} \in H_e^*(X; Z_p)$, *where* β_* *is the ordinary homology Bockstein and* β^* *is the equivariant cohomology Bockstein.*

The proof of (4.16) will be given in § 9 (9.2).

5. The operations $\mathcal{P}^i$ and local coefficients

5.1. Let $\bar{u} \in H_e^q(X; G)$, then following Steenrod [19] set:

$$\mathcal{P}^i\bar{u} = (-1)^{q+i+mq(q+1)/2}(m!)^{2i-q} w_{(q-2i)(p-1)}(\bar{u}) \tag{5.2}$$

where $m = p - \frac{1}{2}$ and $w_{(q-2i)(p-1)}$ is a generator of $H_{(q-2i)(p-1)}(\pi; Z_p)$.
Thus:

$$\mathcal{P}^i \colon H_e^q(X; G) \to H_e^{q+2i(p-1)}(X; G) \tag{5.3}$$

5.4. Let X be a space with base point x_0, $\tilde{X}$ its universal covering space at x_0, then by (3.1)

$$p^* \colon H^*(X; \mathcal{G}) \approx H_e^*(\tilde{X}; G_0).$$

We define the operations $\mathcal{P}^i$ in the local coefficients $\mathcal{G}$ (with basic group a cyclic group of order $p > 2$) by

$$\mathcal{P}^i u = p^{*-1}\mathcal{P}^i(p^* u). \tag{5.5}$$

Thus: $\mathcal{P}^i \colon H^q(X; \mathcal{G}) \to H^{q+2i(p-1)}(X; \mathcal{G})$ and they can be extended to the relative case just as the simple $\mathcal{P}^i$. We now state the elementary properties that the reduced powers $\mathcal{P}^i$ with local coefficients satisfy:

5.6. THEOREM. *The Steenrod powers* $\{\mathcal{P}^i\}$ *of* (5.5) *are a sequence of functions from* $H^*(X; \mathcal{Z}_p)$ *into itself with the following properties:*

1) *They are natural mappings*

$$\mathcal{P}^i \colon H^q(X; \mathcal{G}) \to H^{q+2i(p-1)}(X; \mathcal{G})$$

for every $\mathcal{G} \in \mathcal{Z}_p(X)$

2) *They are homomorphisms*

3) $\mathcal{P}^0$ *is the identity homomorphism. If* $u \in H^q(X; \mathcal{G})$, *then* $\mathcal{P}^s u = 0$ *if* $s > q/2$ *and if* q *is even* $\mathcal{P}^{q/2}u = u^p$ *in the sense of cup products.*[3]

4) *If* $u \in H^q(X; \mathcal{G}_1)$, $v \in H^r(X; \mathcal{G}_2)$, *then* $u \cup v \in H^{q+r}(X; \mathcal{G}_1 \circ \mathcal{G}_2)$ *and* $\mathcal{P}^i(u \cup v) = \sum_{a+b=i} \mathcal{P}^a u \cup \mathcal{P}^b v$

---

[3] If $\mathcal{G} \in \mathcal{Z}_p(X)$, then the p-fold cup product of a class u of $H^*(X; \mathcal{G})$ again lies in $H^*(X; \mathcal{G})$.

5) *If Y is a closed subspace of X, $g\colon Y \to X$ the inclusion and $\delta\colon H^q(Y, g^*\mathcal{G}) \to H^{q+1}(X, Y; \mathcal{G})$ is the coboundary operator of the cohomology sequence of the pair (X, Y), then*

$$\mathcal{P}^i\delta = \delta\mathcal{P}^i$$

6) *$\mathcal{P}^i$ restricted to a simple system of coefficients can be naturally identified with the reduced powers of Steenrod.*

(1) follows easily from (5.3), (5.5), (3.1) and (4.9); (2) to (5) will be proven in § 9 (9.3).

As to (6) we need to remark only that when $\mathcal{G}$ is a simple system of groups, (4.3) when applied to

$$p\colon (\tilde{X}, \pi_1(X, x_0)) \to (X, e)$$

where e denotes the group of one element, gives us a commutative diagram.

$$\begin{array}{ccc} H^q(X; G_0) & \xrightarrow{\mathcal{P}^i} & H^{q+2i(p-1)}(X; G_0) \\ \Big\downarrow p^* & & \Big\downarrow p^* \\ H_e{}^q(\tilde{X}; G_0) & \xrightarrow{\mathcal{P}^i} & H_e{}^{q+2i(p-1)}(\tilde{X}; G_0) \end{array}$$

where in the upper line, $\mathcal{P}^i$ is the reduced powers of Steenrod with ordinary coefficients.

(5.6)-(6) and (3.7) will allow us to compute reduced powers with local coefficients in terms of reduced powers in the universal covering space.

6. Example of spaces which are distinguished by $\mathcal{P}^i$ with non-simple coefficients. We construct two spaces M, N so that no primary operation with ordinary coefficients distinguishes them as to homotopy type, but which a $\mathcal{P}^i$ with local coefficients distinguishes. We only show that $\mathcal{P}^i$ with local coefficients distinguishes them, and leave it as an exercise to the reader to verify that no primary operation does.

Let K denote the complex projective three space (real dimension 6), L the union of the complex projective plane with a 6-sphere with a single point in common, so chosen that it has real coordinates in the projective plane and lies in the equator of the sphere.

Let σ denote the cyclic group of order 2, and let the non-trivial element of σ operate in K by complex conjugation. In L let it operate by complex conjugation in the projective plane, and by reflection along the equator on

the sphere. Let sK and sL denote the suspensions of K and L. The operations of σ extend by suspending the mappings.

Let M be the identification space of (unit interval) $\times sK$ under the relation $(0,a) \sim (1,ga)$ where $a \in sK$, g is the operation in sK of the non-trivial element of σ. Construct N similarly, using sL.

Let I_3 denote the twisted integers mod 3, i.e. the integers mod 3 on which σ acts by change of sign.

Now we have:

6.1. *The operation* $\mathcal{P}^1\colon H^3(\ ;I_3) \to H^7(\ ;I_3)$ *distinguishes* M *and* N.

Notice first that by construction M and N have the same six-skeleton also $\pi_1(M) \approx \pi_1(N) \approx Z$ and if $\eta\colon Z \to \sigma$ is the natural projection, we may use it to define local systems of groups I_3 on M and N.

Let A be the unit interval, then: $\tilde{M} \cong A \times sK$, $\tilde{N} = A \times sL$ and in both cases one can easily check using (3.1) the following isomorphisms:

$$p_1^*\colon H^*(M;I_3) \approx H_e^*(A \times sK;Z_3)$$
$$p_2^*\colon H^*(N;I_3) \approx H_e^*(A \times sL:Z_3)$$

where the equivariant cohomology is taken with respect to the action of σ. If we apply (3.6), we obtain natural isomorphisms:

$$H^*(M;I_3) \approx H^*(A \times sK;Z_3)^\sigma \approx H^*(sK;Z_3)^\sigma$$
$$H^*(N;I_3) \approx H^*(A \times sL;Z_3)^\sigma \approx H^*(sL;Z_3)^\sigma.$$

Now it is easily seen that $H^q(sK;Z_3) = Z_3$ for $q = 0,3,5,7$ and also $H^q(sL;Z_3) = Z_3$ for $q = 0,3,5,7$. If u_1 is a generator of $H^3(sK;Z_3)$, u_2 the corresponding generator in $H^3(sL;Z_3)$ then both u_1 and u_2 are invariant under σ, hence $\mathcal{P}^1 u_1$ and $\mathcal{P}^1 u_2$ are also invariant. Now $\mathcal{P}^1 u_1 \neq 0$, $\mathcal{P}^1 u_2 = 0$ for $u_1 = s^*\bar{u}_1$, $u_2 = s^*\bar{u}_2$ where s^* is the cohomology suspension and it is clear from the construction that $\bar{u}_1{}^3 \neq 0$, $\bar{u}_2{}^3 = 0$. Hence

$$\mathcal{P}^1 u_1 = \mathcal{P}^1 s^* \bar{u}_1 = s^* \mathcal{P}^1 \bar{u}_1 \neq 0,$$

for $\mathcal{P}^1 \bar{u}_1 = \bar{u}_1{}^3$ and $\mathcal{P}^1 u_2 = 0$.

If we take the corresponding elements to u_1 and u_2 in the cohomology with coefficients I_3, the above inequalities hold.

Example of spaces for which $\mathcal{P}^i$ for $i > 1$ distinguishes them as to homotopy type are obtained by taking higher dimensional projective spaces.

6.2. We may remark that the $\mathcal{P}^i$ with non-simple coefficients G (cyclic of order p) do not give any information on H-spaces or homogeneous spaces,

for one can prove that the cohomology with (non-simple) local coefficients in G of such a space is totally acyclic.

This tells us that to find non-trivial applications of $\mathcal{P}^i$ with non-simple coefficients Z_p, one should look at spaces which are far from being H-spaces.

Chapter III.

Cohomology operations with local coefficients.

7. The complex $L_\pi(G, q)$. It universal character.

7.1. In this paragraph we define the general concept of cohomology operations with local coefficients; this is done in the category of Kan complexes (cf. [12]).

By a complex, we will mean a Kan complex.

Let X be a complex, we will say it is one-vertexed if the 0-simplices consist of a single element x_0, and by $\pi_1(X)$ we denote the fundamental group of the one-vertexed complex. If G is an abelian group having $\pi_1(X)$ as group of operators, then G is a local group at x_0 in the sense of Section 1, and we may then speak of the cohomology of X with local coefficients in G. To stress the way in which $\pi_1(X)$ operates in G, let $h\colon \pi_1(X) \to A(G)$ be the mapping defining G as a $\pi_1(X)$-module; then $H^*(X; h, G)$ will denote the cohomology groups of X with local coefficients in G.

7.2. Let π be a multiplicative group, by $\mathcal{L}_\pi$ we denote the category whose objects are pairs (X, α), where X is a one-vertexed complex and α a homomorphism $\pi_1(X) \to \pi$.

The mappings $f\colon (X, \alpha) \to (Y, \beta)$ of $\mathcal{L}_\pi$ are semisimplicial mappings $f\colon X \to Y$ with the restriction that they should induce commutative diagrams:

$$\begin{tikzcd} \pi_1(X) \arrow[rr, "f_*"] \arrow[dr, "\alpha"'] & & \pi_1(Y) \arrow[dl, "\beta"] \\ & \pi & \end{tikzcd} \tag{7.3}$$

Let G be a group having π as a group of (left) operators, that is we are given $h\colon \pi \to A(G)$. Then for every $(X, \alpha) \in \mathcal{L}_\pi$, we have a local group $(h\alpha, G)$, and if $f\colon (X, \alpha) \to (Y, \beta)$ is a mapping, then f induces

$$f^*\colon H^*(Y; h\beta, G) \to H^*(X; h\alpha, G). \tag{7.4}$$

7.5. Let G, G' be abelian groups, $h: \pi \to A(G)$, $h': \pi \to A(G')$ be mappings defining the module structures over π, then by $\mathcal{O}_\pi(h, G, q; h', g', n)$ we denote the set of cohomology operations:

$$T: H^q(\ ; h, G) \to H^n(\ ; h', G') \text{ where for every } (X, \alpha) \in L_\pi,$$

$T(X, \alpha)$ is a function:

$$T(X, \alpha): H^q(X; h\alpha, G) \to H^n(X; h'\alpha, G')$$

which is natural with respect to mappings $f: (X, \alpha) \to (Y, \beta)$, that is, satisfy:

$$f^* T(Y, \beta) = T(X, \alpha) f^*$$

7.6. In Chapter I we reduced the study of cohomology with local coefficients to the study of equivariant cohomology in the covering space.

We define now the semisimplicial analogue of universal covering space following V. K. Guggenheim (cf. [9]).

Let X be a one-vertexed complex and let $\pi = \pi_1(X)$. Then we form a new complex $\tilde{X}$ so that:

$$\tilde{X}_n = \pi \times X_n$$

with face operators

$$\partial_i(\gamma, x) = (\gamma, \partial_i x) \text{ for } 0 < i \leqq n = \dim x$$

and $\partial_0(\gamma, x) = (\gamma w_x, \partial_0 x)$ where w_x is the class in π corresponding to the leading edge of x, i.e. the edge $\partial_2 \partial_3 \cdots \partial_n x$; and degeneracy operators:

$$s_i(\gamma, x) = (\gamma, s_i x) \quad 0 \leqq i \leqq n.$$

To verify that $\tilde{X}$ is a complex which satisfies the usual properties of a universal covering space one uses the description Kan has given (cf. [10]) of the fundamental group of a one-vertexed Kan complex.

If $f: X \to Y$ is a mapping of complexes, then f induces:

$$\tilde{f}: \tilde{X} \to \tilde{Y}$$

by

$$\tilde{f}(\gamma, x) = (f_* \gamma, fx)$$

where $f_*: \pi_1(X) \to \pi_1(Y)$ is the homomorphism of fundamental groups induced by f. $\pi_1(X)$ operates in $\tilde{X}$ freely by multiplication on the first factor, and if $f: X \to Y$ is a mapping then $\tilde{f}: (\tilde{X}, \pi_1(X)) \to (\tilde{Y}, \pi_1(Y))$ is a mapping in the sense of (4.3).

The Eilenberg Theorem (3.1) holds in the category of one-vertexed complexes, and we restate it for convenience:

7.7. *Let X be a one-vertexed complex, π its fundamental groups (h, G) a local group at x_0, then the projection $p: \tilde{X} \to X$, defined by $p(\gamma, x) = x$ induces a natural isomorphism:*

$$p^{\mathfrak{c}}: H^*(X; h, G) \to H_e^*(\tilde{X}; G)$$

7.8. We recall that the cartesian product of two complexes X and Y is the complex $X \times Y$ whose n-simplices are ordered pairs of n-simplices, one of X and one of Y; the face and degeneracy operators are the face and degeneracy operators applied to each factor. We want to construct a sequence of one-vertexed complexes associated with a multiplicative group π and a π-module G. We proceed as follows:

Let $W(\pi)$ be the standard free acyclic complex corresponding to the group π, i.e. the n-simplices of $W(\pi)$ are ordered $(n+1)$-tuples of elements of π, with face and degeneracy operators:

$$\partial_i(a_0, \cdots, a_n) = (a_0, a_2, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n)$$

and

$$s_i(a_0, \cdots, a_n) = (a_0, \cdots, a_i, 1, a_{i+1}, \cdots a_n)$$

where 1 is the unit of π and π operates in $W(\pi)$ by diagonal multiplication.

Let $K(G, q)$ be the standard Eilenberg-MacLane complex (cf. [7]). $K(G, q)$ has a single non-vanishing homotopy group which is G in dimension q. It is one-vertexed. The n-simplices of $K(G, q)$ are in 1-1 correspondence with $Z^q(\Delta_n; G)$ the group of normalized q-cocycles of the standard n-simplex with coefficients in G. The face operators are induced by simplicial mappings $\Delta_{n-1} \to \Delta_n$ which imbed Δ_{n-1} as face of Δ_n, the degeneracy operators induced by the projection: $\mu_i: \Delta_{n+1} \to \Delta_n$, where:

$$\mu_i(j) = j \text{ if } j \leqq i \text{ and } \mu_i(j) = j - 1 \text{ if } j > i.$$

Let $W(\pi) \times K(G, q)$ be the cartesian product of $W(\pi)$ and $K(G, q)$. Since π operates in G, it operates as semisimplicial automorphisms of $K(G, q)$ and we let π operate in $W(\pi) \times K(G, q)$ diagonally, then $L_\pi(G, q)$ will denote the complex obtained from $W(\pi) \times K(G, q)$ by identification under π. $L_\pi(G, q)$ is one-vertexed.

7.9. $L_\pi(G, q)$ is called the (π, q)-universal complex in dimension q. $W(\pi) \times K(G, q)$ is π-equivariantly isomorphic to $\tilde{L}_\pi(G, q)$ under the mapping:

$$\theta: W(\pi) \times K(G, q) \to \tilde{L}_\pi(G, q) \tag{7.10}$$

given by:

$$(7.11) \qquad \theta((a_0,\cdots,a_n),x_n) = (a_0,[a_0,a_1,\cdots,a_n,x_n])$$

where $[a_0,\cdots,a_n,x_n]$ is the class of $((a_0,\cdots,a_n),x_n)$ in $L_\pi(G,q)$.

By (7.10), when we speak of the universal covering of $L_\pi(G,q)$, we will have in mind $W(\pi)\times K(G,q)$.

7.12. We now show that $L_\pi(G,q)$ has a canonical class in

$$H^q(L_\pi(G,q);h,G)$$

where $h\colon \pi\to A(G)$ is the mapping making G a π-module. Consider then:

$$(7.13) \qquad L_\pi(G,q) \xleftarrow{\ p\ } W(\pi)\times K(G,q) \xrightarrow{\ p_2\ } K(G,q)$$

where p_2 is the projection onto the second factor. It is equivariant so (7.13) induces:

$$H^q(L_\pi(G,q);h,G) \xrightarrow{\ p^*\ } H_e{}^q(W(\pi)\times K(G,q)) \xleftarrow{\ p_2{}^*\ } H_e{}^q(K(G,q);G)$$

The fundamental class (cf. [4]) of $H^q(K(G,q);G)$ is in fact equivariant. Let us denote it by u_0 then we define

$$U_0\in H^q(L_\pi(G,q);h,G)$$

by

$$(7.14) \qquad U_0=(p^*)^{-1}p_2{}^*u_0.$$

We define now $i\colon K(G,p)\to L_\pi(G,q)$ by

$$(7.15) \qquad i(x_n)=[(s_0{}^n)1,x_n]$$

where $x_n\in K(G,q)_n$ and $(s_0{}^n)1$ is the n-th fold iteration of the 0-th degeneracy applied to the unit of π. i is one-one into and will be referred to as the inclusion of $K(G,q)$ in $L_\pi(G,q)$. For any local group (h',G'), i induces:

$$i^*\colon H^*(L_\pi(G,q);h',G')\to H^*(K(G,q);G')$$

Now it follows that

$$(7.16) \qquad i^*U_0=u_0.$$

7.17. Let X, Y be one-vertexed complexes, α a homomorphism: $\pi_1(X)\to\pi_1(Y)$ then by $\pi_\alpha(X,Y)$ we denote the set of homotopy classes of mappings with base point fixed, which induce α on the fundamental groups.

We have now a classification theorem:

7.18. THEOREM. *Let* $(X,\alpha) \in \mathcal{L}_\pi$, *then the function which to every* $\phi \in \pi_\alpha(X; L_\pi(G,q))$ *assigns* $\phi^* U_0$ *in* $H^q(X; h\alpha, G)$ *establishes a one-to-one corerspondence between*

$$\pi_\alpha(X, L_\pi(G,q))$$

and

$$H^q(X; h\alpha, G).$$

(7.18) is a special case of a theorem of P. Olum [14]. It permits us to deduce just as in the case of ordinary coefficients the following:

7.19 THEOREM. *The function which to every* $T \in \mathcal{O}_\pi(h, G, q; h', G', n)$ *associates* $T(L_\pi(G,q), 1) U_0 \in H^n(L(G,q); h', G')$ *establishes a one-to-one corerspondence between* $\mathcal{O}_\pi(h, G, q; h', G', n)$ *and* $H^n(L_\pi(G,q); h', G')$.

The role of π to link the groups G and G' as far as cohomology operations are concerned is made clearer by the following:

7.20 THEOREM. *Let* $u \in H^n(L_\pi(G,q); h', G')$ *come from* $H^n(K(\pi,1); h', G')$ *under the projection* $L_\pi(G,q) \to K(\pi,1)$. *Then the cohomology operation* T_u *uniquely associated to* u *by* (7.19) *is constant, i.e. for any* $(X,\alpha) \in \mathcal{L}_\pi$ *and any pair of classes* $x, y \in H^q(X; h\alpha, G)$

$$T_u(x) = T_u(y).$$

The proofs of (7.18), (7.19) and (7.20) will be given in Section 10.

7.21. In Chapter II we defined reduced powers with local coefficients and singular theory. We may extend the definition of the $\mathcal{P}^i$ to Kan complexes as follows:

If X is a Kan complex, let $|X|$ denote the geometric realization of X in the sense of Milnor [11] and $S|X|$ the total singular complex of $|X|$. Milnor defines a homotopy equivalence:

$$k_X : X \to S|X|$$

which is functorial with respect to s.s. mappings. Choose a base point x_0 in X, then $k_{X*} : \pi_1(X, x_0) \approx \pi_1(S|X|, k_X(x_0))$, and if G_0 is a local group at x_0, we may consider G_0 as a local group at $k_X(x_0)$ and then k_X induces:

$$k_X^* : H^*(S|X|; G_0) \to H^*(X; G_0)$$

If now X is an one-vertexed (Kan) complex, G a local group, at the base point, which is cyclic of order p; define

$$\mathcal{P}^i : H^q(X;G) \to H^{q+2i(p-1)}(X;G)$$

by:

(7.22) $$\mathcal{P}^i u = f_X^{*-1}\mathcal{P}^i k_X^* u$$

then it easily follows that the $\mathcal{P}^i$ satisfy the properties of (5.6) and we have:

7.23. *The reduced powers* $\{\mathcal{P}^i\}$ *are a sequence of cohomology operations defined in the category* $\mathcal{L}_{A(Z_p)}$. *For every pair of integers i and q and any cyclic group G of order p, they give rise to an operation* $\mathcal{P}^i \in \mathcal{O}_{A(Z_p)}(G,q;G,q+2i(p-1))$ *where* $A(Z_p)$ *operates in G via a fixed isomorphism* $G \approx Z_p$.

8. Operations with initial and terminal (local) coefficients cyclic groups of order p.

8.1. *Notations.* In this section π will denote the automorphism group $A(Z_p)$ of the integers mod p, and G will denote Z_p; G' will denote a cyclic group of order p (p-prime > 2) and $\mathcal{O}_\pi(G,q;G',n)$ will denote the set of cohomology operations with initial coefficient group G on which π operates with the standard action, while G' is the terminal coefficient group on which π acts via a fixed isomorphism $G' \to G$.

For every space X we have defined a group $\mathcal{G}_p(X)$ (see § 2) of local coefficient systems with basic group cyclic of order p, and $H^*(X;\mathcal{G}_p)$ stands for the bigraded vector space over Z_p with bidegrees $(q,\mathcal{B}')$ where q is a non-negative integer and $\mathcal{B}' \in \mathcal{G}_p(X)$.

The main result in this section (see (8.5)) asserts that $H^*(L_\pi(G,q);\mathcal{G}_p)$ is naturally isomorphic to $H^*(K(G,q);G)$. Cartan [4] has given a complete description of the cohomology of $H^*(K(G,q);G)$. Easy consequences of this are then (5.6), a characterization of reduced powers with local coefficients, and a system of generators for operations with initial and terminal groups G and G' respectively.

8.2. We begin by recalling the structure of $H^*(K(G,q);G)$ as given by Cartan. A sequence $I = (a_1, \cdots, a_k)$ of integers is called an admissible sequence of height q and stable degree $q(I)$ if:

1) $a_i \equiv 0, 1 \bmod 2(p-1)$
2) $a_i \geqq p a_{i+1}$

3) $\sum a_i = q(I)$

and

4) $pa_1 < (p-1)(q+q(I))$.

To every admissible sequence I of height q we associate an operation St^I as follows:

By (1) $a_i = 2k_i(p-1) + e_i$ where $e_i = 0, 1$. If $e_i = 0$, set $St^{a_i} = \mathfrak{P}^{k_i}$, and if $e_i = 1$ set $St^{a_i} = \mathfrak{P}^{k_i}\beta$. Then

$$St^I = St^{a_1} St^{a_2} \cdots St^{a_k}$$

corresponds to the admissible sequence I.

8.3 THEOREM (Cartan). *The cohomology ring $H^*(K(G,q);G)$ is a tensor product of exterior and polynomial algebras on one generator. The generators of these algebras are $St^I u_0$, where u_0 is the fundamental class of $K(G,q)$ and I ranges over all admissible sequences of height q. If $St^I u_0$ is even (odd) dimensional, it generates a polynomial (exterior) algebra.*

The proof of (8.3) can be found in [4]. The system of algebra generators in (8.3) will be referred to as the *Cartan generators* of $H^*(K(G,q);G)$.

8.4. Now let us look at $L_\pi(G,q)$. $\mathcal{G}_p(L_\pi(G,q))$ is cyclic of order $p-1$, and a generator for this group is the class $\mathcal{G}$ of local groups with representative the local group G and standard action of π on G. Let $\mathcal{G}^j$ be the j-fold product of $\mathcal{G}$.

Since $K(G,q)$ is simply connected if $q > 1$, $\mathcal{G}_p(K(G,q))$ reduces to G, and the inclusion:

$$i \colon K(G,q) \to L_\pi(G,q)$$

of (7.15) induces by (2.5) a ring homomorphism:

$$i^* \colon H^*(L_\pi(G,q);\mathcal{G}_p) \to H^*(K(G,q);G)$$

with $i^* U_0 = u_0$, where U_0 is the fundamental class of $L_\pi(G,q)$ and u_0 that of $K(G,q)$.

By the *degree of a monomial* on the Cartan generators we will mean the number of factors in the monomial.

Now we can state our main theorem:

8.5. *Let* $i\colon K(G,q)\to L_\pi(G,q)$ *be the inclusion, then* i *induces a ring isomorphism:*

$$i^*\colon H^*(L_\pi(G,q)\,;\mathfrak{F}_p)\to H^*(K(G,q)\,;G)$$

with

$$i^*U_0=u_0.$$

Furthermore for every j, i^* *sends* $H^*(L_\pi(G,q)\,;\mathfrak{G}^j)$ *isomorphically onto a vector subspace* V_j *of* $H^*(K(G,q)\,;G)$; *a basis for* V_j *consists of all those monomials in the Cartan generators of degree* k, *where* $k\equiv j \bmod p-1$.
A consequence of (8.3) and (8.5) is:

8.6. *The set of cohomology operations in* $\mathcal{L}_\pi$ *with initial group* G, *and terminal group* G' (*then action of* π *in* G' *being arbitrary*) *has a system of generators which consists of: addition, cup products, Bockstein homomorphisms and reduced powers* $\mathfrak{P}^i$.

We also have:

8.7. Uniqueness Theorem. *There is only one set of cohomology operations with initial and terminal local systems with basic group a cyclic group of order* p, *which reduces to the ordinary* $\mathfrak{P}^i$.

Chapter IV.

The Proofs.

9. The proofs of Chapter II.

9.1. *The proof of* (4.2) *and* (4.3). As we mentioned earlier we use the acyclic models of Eilenberg-MacLane. For the main theorem and concepts we refer to [8].

Let $\mathfrak{A}$ denote the category whose objects are pairs (X,ρ) as in (4.3). Let $\partial\mathfrak{G}_*$ denote category whose objects are pairs (A,ρ) where A is a (free) chain complex, ρ a group acting as automorphisms of A. A mapping $g\colon(A,\rho)\to(B,\sigma)$ is a chain mapping $A\to B$ and a homomorphism $\rho\to\sigma$ both denoted by g such that:

$$g(sa)=g(s)g(a)$$

We consider the following models $\mathfrak{M}$ in $\mathfrak{A}$, $(\Delta_q,1)$ $q=0,1,\cdots$ where 1 is the group of one element.

Let π be a permutation group on p symbols (p a prime > 2) and W a fixed π-free acyclic chain complex.

Define two functors K, L: $\mathfrak{A} \to \partial\mathfrak{G}_*$ as follows:

$$(9.2) \qquad K(X,\rho) = (W \otimes C_*(X), \pi \times \rho)$$

where $C_*(X)$ are the integral singular chains on X, and $\pi \times \rho$ acts in $W \otimes C_*(X)$ diagonally.

If $f: (X,\rho) \to (Y,\sigma)$ then $F(f): K(X,\rho) \to K(Y,\sigma)$ is the pair $(1 \otimes f_\#, 1 \times f)$.

Now

$$(9.3) \qquad L(X,\rho) = ([C_*(X)]^p, \pi \times \rho)$$

where $[C_*(X)]^p$ is the p-fold tensor product of $C_*(X)$ with itself and $\pi \times \rho$ operates in $[C_*(X)]^p$ by:

$$(r,s)(x_1 \otimes \cdots \otimes x_p) = \epsilon_r (sx_1) \otimes \cdots \otimes (sx_{i+1}) \otimes (sx_i) \otimes \cdots \otimes (sx_p)$$

where $\epsilon_r = (-1)^{\dim x_i \dim x_{i+1}}$ if r is the simple interchange of the i-th and $(i+1)$ factors and s is any element of ρ.

If f: $(X,\rho) \to (Y,\sigma)$, then $L(f): L(X,\rho) \to L(Y,\sigma)$ is the pair $((f_\#)^p, 1 \times f)$.

Let $\mathfrak{G}_*$ denote the category of pairs (A,ρ) where ρ is a group, A a ρ-module. Define

$$K_{p,q}: \mathfrak{A} \to \mathfrak{G}_*$$

by

$$K_{p,q}(X,\rho) = (W_p \otimes C_q(X), \pi \times \rho)$$

and an associated functor (cf. [8]):

$$\hat{K}_{p,q}: \mathfrak{A} \to \mathfrak{G}_*$$

where $\hat{K}_{p,q}(X,\rho)$ is the $\pi \times \rho$-free module generated by pairs (u,m) where $u: (\Delta_q, 1) \to (X,\rho)$ is a mapping in $\mathfrak{A}$ and $m \in K_{p,q}(\Delta_q, 1)$. We have a natural transformation of functors:

$$\Phi: K_{p,q} \to K_{p,q}$$

defined by $\Phi(X,\rho)(u,m) = K_{p,q}(u)(m)$.

Construct now:

$$\Xi\colon \hat{K}_{p,q} \to K_{p,q}$$

by

$$\Xi(X,\rho)(w \otimes x) = (w \otimes \Delta_q, u)$$

defined in the $\pi \times \rho$-basis of $K_{p,q}(X)$. $u\colon (\Delta_q, 1) \to (X,\rho)$ is the unique mapping taking the q-chain Δ_q into x.

It is easy to see that Ξ is a natural transformation and that $\Phi X =$ identity: $\Xi_{p,q} \to K_{p,q}$. Define now: $K_p = \sum_{r+s=p} K_{r,s}$, $\hat{K}_p = \sum_{r+s=p} \hat{K}_{r,s}$ and extend Φ and Ξ.

The functor L is acyclic on models. For we can produce a contracting homotopy

$$s\colon C_n(\Delta_q) \to C_{n+1}(\Delta_q)$$

and obtain a contracting homotopy for $[C_*(\Delta_q)]^n$, which we will still denote by s.

Let $L_q\colon A \to G^*$ be defined by $L_q(X,\rho) = \{x \in (C_*(X))^p$ of dimension $q\}$. Then:

$$\partial s + s\partial = id\colon L_n(\Delta_q, 1) \to L_n(\Delta_q, 1)$$

Now consider:

$$h_0\colon K_0 \to L_0$$

defined by $h_0(X,\rho)(w_0 \otimes x_0) = x_0 \otimes \cdots \otimes x_0$ if w_0, x_0 are vertices of W and X, and extend h_0 by linearity to $K_0(X,\rho)$.

We may now apply Theorem II of [8] to obtain an extension of h_0 which is the required ϕ' of (4.2) which satisfies (4.3).

9.2. *Proof of* (4.16).

By (4.15) the operation β^* is natural, hence reduces to the ordinary Bockstein when the coefficient system is simple. Also it is a cohomology operation in $\mathcal{O}_{A(Z_p)}(Z_p, q, Z_p, q+1)$, hence to prove (4.16) it suffices to do it in the universal example. With the notation of (8.5) and (4.16):

$$i^*((\beta_*\xi)U_0 - \beta^*(\xi(U_0)) = (\beta_*\xi)u_0 - \beta^*(\xi(u_0)).$$

Now in [19] it is shown that the right hand side is 0.

Then (4.16) follows by (8.5).

9.3. *Proof of* (5.6), *parts* (2) *to* (5).

i) (2) to (4) are proved in the same way as (4.16), namely the $\mathcal{P}^i$ are natural by (5.6) (1), hence it suffices to prove the assertions for the uni-

versal example. In [19] it is shown that the corersponding formulas hold for simple coefficients. Then (8.5) shows the same is true for local coefficients.

ii) We prove (5) directly:

Let (X, Y) be a pair, $\mathcal{B} \in \mathfrak{F}_p(X)$, $i\colon Y \to X$ and $j\colon X \to (X, Y)$ the inclusions then we have an exact sequence:

(9.4) $$H^q(Y; i^*\mathcal{B}) \longrightarrow H^{q+1}(X, Y; \mathcal{B}) \xrightarrow{j^*} H^{q+1}(X; \mathcal{B}) \xrightarrow{i^*} H^{q+1}(Y; \mathcal{B})$$

and we want to prove:

(9.5) $$\mathcal{P}^i \delta u = \delta \mathcal{P}^i u$$

for every $u \in H^q(Y; i^*\mathcal{B})$.

We reduce the proof of (9.5) by a sequence of transformations:

$$(X, Y) \xrightarrow{f_1} (X, Y \times I) \xleftarrow{f_2} (\bar{X}, Y \times 1)$$

$$(\bar{X}, Y \times 1) \xrightarrow{f_3} (\bar{X}, X \cup (Y \times 1)) \xleftarrow{f_4} (Y \times I, Y \times \dot{I})$$

where we need to define the new terms and mappings. I is the unit interval $\dot{I} = \{0,1\}$ its end points and $\bar{X} = X \cup (Y \times I)$. f_1 is the inclusion; it induces a natural isomorphism of the exact sequences (9.4). f_2 is the inclusion and since $Y \times 1$ is a deformation retract of $Y \times I$, f_2 induces an isomorphism of the cohomology sequences (9.4). f_3 is the inclusion and it induces an epimorphism:

(9.6) $$f_3^*\colon H^*(X \cup (Y \times 1); \mathcal{B}) \to H^*(Y \times 1; \mathcal{B}).$$

Finally f_4 is an excision, hence also induces an isomorphism of the exact sequence (9.4).

By naturality of the $\mathcal{P}^i$ it suffices to prove (9.5) for the pair $(\bar{X}, Y \times 1)$. Assume (9.5) holds for the pair $(Y \times I, Y \times \dot{I})$, then it also holds for $(\bar{X}, X \cup (Y \times 1))$. Since the $\mathcal{P}^i$ are natural, δ commutes with f_3^*, and f_3^* in (9.6) is onto, one readily proves that the $\mathcal{P}^i$ commute with δ of the pair $(\bar{X}, Y \times 1)$. Since f_2^* and f_1^* are isomorphisms, it follows that (9.5) holds for (X, Y).

Hence it remains to see that (9.5) holds for the pair $(Y \times I, Y \times \dot{I})$. Let $\mathcal{B} \in \mathfrak{F}_p(X)$, $\mathcal{B}'$ the induced system in $Y \times I$ and consider

$$H^q(Y\times \dot I;\mathcal{B}') \xrightarrow{\ \delta\ } H^{q+1}(Y\times I, Y\times \dot I;\mathcal{B}')$$

$H^q(Y\times \dot I)$ consists of two copies of $H^q(Y)$, obtained by crossing with the generator of the 0-cohomology groups of the end points of I. Let $\bar a_0$ and $\bar a_1$, be the respective generators. Thus if $u\in H^q(Y)$, u gives rise to two elements in $H^q(Y\times \dot I)$, $u\times \bar a_0$ and $u\times \bar a_1$. Let I' denote the generator of $H^1(I,\dot I;Z_p)$, then if $u\in H^q(Y;\mathcal{B})$, $u\times I'\in H^{q+1}(Y\times I, Y\times \dot I;\mathcal{B}')$ and it is shown in [13] that:

$$\delta(u\times \bar a_i) = \epsilon(u,i)u\times I'$$

where $\epsilon(u,i) = (-1)^{\dim u+i}$. Hence

$$\begin{aligned}\delta\mathcal{P}^i(u\times \bar a_i) &= \delta(\mathcal{P}^i u\times \bar a_i) = \epsilon(u,i)\mathcal{P}^i u\times I'\\ &= \epsilon(u,i)\mathcal{P}^i(u\times I') = \mathcal{P}^i\delta(u\times \bar a_i)\end{aligned}$$

by parts (2), (3) of (5.6) together with formula (4) for cross-products.

10. The proofs of Chapter III.

10.1. *Proof of* (7.18). We sketch a proof of (7.18) by passing to the universal covering spaces. Consider then the diagram:

(10.2)
$$\begin{tikzcd}
\pi_\alpha(X, L_\pi(G,q)) \arrow[r, "\nu_1"] \arrow[dd, "\eta_1"] & \pi_\alpha(\tilde X, W(\pi)\times K(G,q)) \arrow[d, "\nu_2"] \\
 & \pi(\tilde X, K(G,q)) \arrow[d, "\eta_2"] \\
H^q(X;h\alpha, G) & H_e^q(X;G) \arrow[l, "p^*"']
\end{tikzcd}$$

1) $\pi_\alpha(X, L_\pi(G,q))$ is the set of homotopy classes of mappings with base point fixed, inducing α on the fundamental groups.

2) $\pi_\alpha(\tilde X, W(\pi)\times K(G,q))$ is the set of homotopy classes of mappings with base point fixed which are α-equivariant, i.e. satisfy

$$fs = \alpha(s)f$$

for any $s\in\pi_1(X)$.

3) $\pi_\alpha(\tilde X, K(G,q))$ is the set of α-equivariant homotopy classes with base point fixed.

The mappings of (10.2) are defined as follows:

4) $\nu_1(f) = \bar{f}$ where $\bar{f}$ is defined in (7.6). ν_1 passes to classes.

5) $\nu_2(\bar{f}) = p_2\bar{f}$, where $p_2\colon W(\pi) \times K(G,q) \to K(G,q)$ is the π-equivariant projection onto the second factor.

6) $\eta_2(\bar{f}) = \bar{f}^* u_0$.

7) $\eta_1(f) = f^* U_0$.

From the definition of u_0 and U_0, it is easily seen that

$$\eta_1 = p^* \eta_2 \nu_2 \nu_1.$$

Now p^* is an isomorphism, and it is not hard to see that ν_1 is a 1-1 correspondence. As to ν_2, we have the following commutative diagram:

$$\begin{tikzcd} \pi_\alpha(\tilde{X}, W(\pi) \times K(G,q)) \arrow[r, "\nu_2"] \arrow[d, "i_1"'] & \pi_\alpha(\tilde{X}, K(G,q)) \arrow[d, "i_2"] \\ \pi(\tilde{X}, W(\pi) \times K(G,q)) \arrow[r, "p_2"] & \pi(\tilde{X}, K(G,q)) \end{tikzcd}$$

where $\pi(\tilde{X}, W(\pi) \times K(G,q))$ and $\pi(X, K(G,q))$ are the sets of homotopy classes of mappings with base point fixed. So i_1 and i_2 are 1-1 into, while p_2 is a 1-1 correspondence, since $W(\pi)$ is contractible over itself holding the baset point fixed. Now ν_2 is onto; for let $[\bar{f}] \in \pi_\alpha(\tilde{X}, K(G,q))$ and let $g_\alpha\colon \tilde{X} \to W(\pi)$ be the unique α-equivariant homotopy class (with base point fixed), then setting

$$f^1 = (g_\alpha, f) \text{ i.e.}$$
$$f^1(x) = (g_\alpha(x), \bar{f}(x))$$

then $f^1 \in \pi_\alpha(\tilde{X}, W(\pi) \times K(G,q))$ and $\nu_2[f^1] = [\bar{f}]$.

Finally η_2 is a 1-1 correspondence, since the proof Cartan [4] gives of the 1-1 correspondence between $\pi(\tilde{X}, K(G,q))$ and $H^q(\tilde{X}; G)$, shows that α-equivariant mappings correspond to α-equivariant cohomology classes.

10.3. *Proof of* (7.19).

Let T be a cohomology operation in $\mathcal{O}_\pi(h, G, p; h', G', n)$, then $T(U_0)$ is an element of $H^n(L_\pi(G,q); h', G')$. Conversely given $u \in H^n(L_\pi(G,q); h', G')$ define a function

$$T_u\colon H^q(\ \ ; h, G) \to H^n(\ \ ; h', G')$$

in $\mathcal{L}_\pi$ as follows: If $x \in H^q(X; h\alpha, G)$, let ϕ_x denote the unique homotopy class in $\pi_\alpha(X, L_\pi(G,q))$ which corresponds to x by (7.18), define $T_u(x) = \phi_x^*(u)$. T_u is natural, for if $f\colon (Y,\beta) \to (X,\alpha)$ is a mapping in $\mathcal{L}_\pi$, $x \in H^q(X, h\alpha; G)$ then $f^*x \in H^q(Y; h\beta, G)$ and if ϕ_x is as above, then $\phi_x f$ corresponds to f^*x, so

$$T_u(f^*x) = f^*\phi_x^*(u) = f^*(T_u(x)).$$

Hence T_u is a cohomology operation. Moreover it is clear that the correspondences $T \to T(U_0)$ and $u \to T_u$ are inverses of each other, but this is (7.19).

10.4. *Proof of* (7.20). The hypothesis of (7.20) give us a diagram

$$\begin{tikzcd}
X \arrow[r, "\phi_x"] & L_\pi(G,q) \arrow[r, "\phi_u"] \arrow[dr, "p_1"'] & L_u(G', n') \\
& & K(\pi, 1) \arrow[u, "\phi_v"']
\end{tikzcd}$$

in which the triangle is commutative up to homotopy. Here ϕ_u and ϕ_v are the homotopy classes corresponding to u and v according to (7.18). The definition of $T_u(x)$ for a class $x \in H^q(X; h\alpha, G)$ of (10.3) uses $\phi_x \in \pi_\alpha(X, L_\pi(G,q))$ but $\{p\phi_x\} = \{p\phi_y\}$ for any $x, y \in H^q(X; h\alpha, G)$, hence (7.20).

10.5. *Proof of* (8.5). Recall that $L_\pi(G,q)$ was obtained from $W(\pi) \times K(G,q)$ by collapsing under the action of π; let

$$p\colon W(\pi) \times K(G,q) \to L_\pi(G,q)$$

be the natural projection and consider the following sequence:

$$\begin{tikzcd}
H^*(L_\pi(G,q); h', G') \arrow[r, "p^*"] & H_e^*(W(\pi) \times K(G,q); G') \arrow[d, "i^*"] \\
H^*(K(G,q); G') \arrow[r, "p_2^*"] & H^*(W(\pi) \times K(G,q); G')
\end{tikzcd}$$

where i^* is induced by the inclusion of cochains, p_2^* is induced by the pro-

jection onto the second factor, p^* is an isomorphism by (7.7). p_2^* is a π-isomorphism, since $W(\pi)$ is acyclic and p_2 is π-equivariant.

Under the hypothesis of (8.5), π is a cyclic group of order $p-1$, G' a cyclic group of order p, hence we may apply (3.6) to conclude that i is a monomorphism, its image being $H^*(W(\pi)\times K(G,q);G')^\pi$, we thus obtain an isomorphism:

$$\beta\colon H^*(L_\pi(G,q);h',G')\to H^*(K(G,q);G')^\pi \tag{10.6}$$

where $\beta=p_2^{*-1}i^*p^*$.

Now as in Section 2 we are interested in the equivalence class of local systems containing (h',G'), hence we may assume that $(h',G')=\mathfrak{g}^j$ (for some value of $j<p-1$) the j-fold product of the generating class of $\mathfrak{F}_p(X)$ whose representative is $G=Z_p$ with standard action of $A(Z_p)=\pi$.

We may take as representative of $\mathfrak{g}^j$, G with action of π defined as follows:

$$s^*g=s^j\cdot g \tag{10.7}$$

where $g\in G$, $s\in\pi$. Thus (10.6) reduces to:

$$\beta_j'\colon H^*(L_\pi(G,q);h,G')\to H^*(K(G,q);G)^\pi \tag{10.8}$$

where the action of π in the coefficient group G is that given in (10.7). Let us say that π operates in $H^*(K(G,q);G)$ *according to* ϕ_j if it operates with the standard action in $K(G,q)$ and according to (10.7) in the coefficient group. Now we have:

10.9. LEMMA. *The subspace V_j of pointwise fixed elements of $H^*(K(G,q);G)$ under the ϕ_j-action of π, has a basis consisting of all monomials in the Cartan generator of degree r, where r is congruent to $j \bmod (p-1)$.*

Proof. For any $s\in\pi$, let $\eta_j(s)$ be the automorphism of $H^*(K(G,q);G)$ induced by the automorphism of the coefficient group G given by:

$$\eta_j(s)\cdot g=s^j\cdot g$$

and let ψ_s be the automorphism of $H^*(K(G,q);G)$ induced by the action of s in $K(G,q)$, but assuming trivial action on the coefficient group.

We now claim:

$$\phi_j(s)x=(\eta_j(s)\psi_s)x \tag{10.10}$$

for any $x\in H^*(K(G,q);G)$.

It suffices to verify (10.10) in the cochain level. Let $u \in C^n(K(G,q);G)$, then:

$$(\phi_j(s)u)(\sigma) = s^j u(s^{-1}\sigma) = (\eta_j(s)\psi_s u)(\sigma)$$

for any n-simplex σ of $K(G,q)$.

Now let T be a generator of π corresponding to a primitive root k of p, i. e. $T\colon G \to G$ is given by $T(1) = k$.

We determine now the automorphism ψ_T of $H^*(K(G,q);G)$ using the structure of $H^*(K(G,q);G)$.

Let z_0 be the fundamental q-cocycle of $K(G,p)$, then $Tz_0 = k^{-1}z_0$, hence passing to cohomology we obtain:

$$Tu_0 = k^{-1}u_0. \tag{10.11}$$

But from (8.3) u_0 generates $H^*(K(G,q);G)$ by cohomology operations; since ψ_T is natural, this enables us to determine ψ_T in $H^*(K(G,q);G)$. One readily verifies that:

$$\psi_T(m_r) = k^{-r}m_r \tag{10.12}$$

for any monomial m_r on the Cartan generators of degree r.

Since $\phi_j(s)$ is a coefficient automorphism, we see that $\phi_j(T)(x) = k^j x$ for any $x \in H^*(K(G,q);G)$. Combining this with (10.12) and (10.10), we obtain:

$$\phi_j(T)(m_r) = k^{j-r}m_r \tag{10.13}$$

and since k is a primitive root of p,

$$\phi_j(T)(m_r) = m_r$$

if and only if $r - j \equiv 0 \bmod (p-1)$.

Thus m_r is pointwise fixed under the ϕ_j-action of π if and only if $r \equiv j \bmod (p-1)$ where $r =$ degree m_r.

Suppose now $x \in H^*(K(G,q);G)$ is pointwise fixed under the ϕ_j-action of π. Relative to the basis of $H^*(K(G,q);G)$ consisting of monomials in the Cartan generators, x has a representation:

$$x = \sum a_i m_{r_i}$$

where m_{r_i} is a monomial in the Cartan generator of degree r_i and dimension that of x. Then:

$$\phi_j(T)(x) = \sum a_i \phi_j(T) m_{r_i} = \sum a_i m_{r_i}$$

but by the uniqueness of the representation of x in terms of the basis, it follows that $\phi_j(T)(m_{r_i}) = m_{r_i}$, and by the above this implies $r_i \equiv j \bmod (p-1)$. This finishes the proof of (10.9).

10.14. COROLLARY. *For* $1 \leqq j \leqq p-1$,

$$\beta_j' : H^*(L_\pi(G,q); \mathfrak{Z}^j) \to H^*(K(G,q); G$$.

(*where* β_j' *is as in* (10.8)) *is a monomorphism onto* V_j; $V_j \cap V_k = 0$ *if and only if* $j \neq k$, *and* $\sum_{j=0}^{p-1} V_j = H^*(K(G,q); G)$.

Now by (2.4) $i^* : H^*(L_\pi(G,q); \mathfrak{Z}_p) \to H^*(K(G,q); G)$ is a ring homomorphism and (10.14) shows it is in fact a ring isomorphism. (10.9) tells us the structure of $H^*(L(G,q); G^j)$ and thus (8.5) is complete.

PRINCETON UNIVERSITY AND
CENTRO DE INVESTIGACIÓN DEL IPN, MÉXICO, D. F.

REFERENCES.

[1] J. Adem, "The relation on Steenrod powers of cohomology classes," *Algebraic Geometry and Topology*, Princeton University Press (1957), pp. 191-238.

[2] S. Araki, "On the Steenrod's reduced powers in singular homology theories," *Memoirs of the Faculty of Sciences, Kyushu University*, vol. 9 (1956), pp. 159-173.

[3] A. Borel and J.-P. Serre, "Groupes de Lie et puissance reduites de Steenrod," *American Journal of Mathematics*, vol. 75 (1953), pp. 409-448.

[4] H. Cartan, "Algébres d'Eilenberg-MacLane et homotopie," *École Normal Superiere*, Paris, 1956.

[5] H. Cartan and E. Eilenberg, *Homological Algebra*, Princeton University Press, 1956.

[6] S. Eilenberg, "Homology of spaces with operators I," *Transactions of the American Mathematical Society*, vol. 61 (1947), pp. 378-417.

[7] S. Eilenberg and S. MacLane, "On the groups $H(\pi, n)$, I," *Annals of Mathematics*, vol. 60 (1954), 49-139.

[8] ———, "Acyclic models," *American Journal of Mathematics*, vol. 75 (1953), pp. 189-199.

[9] V. K. A. M. Guggenheim, "On a theorem of E. H. Brown," *Illinois Mathematical Journal*, vol. 4 (1960).

[10] D. M. Kan, "Abstract homotopy I," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 41 (1955), pp. 1092-1096.

[11] J. Milnor, "The geometric realization of a semisimplicial complex," *Annals of Mathematics*, vol. 65 (1957), pp. 357-362.

[12] J. C. Moore, "Semisimplicial complexes and Postnikov systems," *Symposium Internacional de Topología Algebraica*, Universidad de México Press (1958), pp. 232-247.

[13] P. Olum, "Obstructions to extensions and homotopies," *Annals of Mathematics*, vol. 52 (1950), pp. 1-50.

[14] ———, "On mappings into spaces in which certain homotopy groups vanish," *Annals of Mathematics*, vol. 57 (1953), pp. 561-573.

[15] K. Reidemeister, *Topologie der Polyeder*, Leipzig, Akad. Verlageges, 1938.

[16] N. E. Steenrod, "Homology with local coefficients," *Annals of Mathematics*, vol. 44 (1943), pp. 610-627.

[17] ———, *Topology of Fibre Bundles*, Princeton University Press, 1951.

[18] ———, "Products of cocycles and extension of mappings," *Annals of Mathematics*, vol. 48 (1947), pp. 290-320.

[19] ———, "Homology groups of symmetric groups and reduced powers operations," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 39 (1953), pp. 213-217.

[20] ———, "Cyclic reduced powers of cohomology classes," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 39 (1953), pp. 217-223.

[21] ———, "Cohomology operations derived from the symmetric group," *Commentarii Mathematici Helvetici*, vol. 31 (1956), pp. 195-218.

[22] E. Thomas, "A generalization of the Pontryagin square cohomology operation," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 42 (1956), pp. 266-269.

ON THE DEFORMATIONS OF LATTICE IN A LIE GROUP.*

By Hsien-Chung Wang.[1]

1. **Introduction.** By a lattice in a topological group G, we shall always mean a discrete subgroup Γ of G with G/Γ compact. Suppose Γ_0 to be a fixed lattice in G, and $r_0\colon \Gamma_0 \to G$ the identity map. Let $\mathcal{R}(\Gamma_0, G)$ denote the space, with the compact-open topology, of all the isomorphisms $r\colon \Gamma_0 \to G$ such that $r(\Gamma_0)$ are lattices in G, and $\mathcal{R}_1(\Gamma_0, G)$ the connected component of $\mathcal{R}(\Gamma_0, G)$ which contains the identity map r_0. Since $\mathcal{R}_1(\Gamma_0, G)$ is arc-wise connected, each element r of $\mathcal{R}_1(\Gamma_0, G)$ will be called a deformation of Γ_0. Those deformations which are obtained by applying automorphisms of G to Γ_0 will be called trivial. When G is a compact or a nilpotent Lie group, it follows directly from some results of Montgomery-Zippin [11] and Malcev [9] that all the deformations of lattices in G are trivial. A. Selberg [13] and Calabi-Vesentini [6] have proved the triviality of deformation of lattices in certain simple Lie groups. Recently A. Weil has established some general properties of the space $\mathcal{R}_1(\Gamma_0, G)$ for an arbitrary Lie group G [14], and also determined completely $\mathcal{R}_1(\Gamma_0, G)$ when G is semi-simple and has no compact factor [15].

Now assume that G is a connected and simply-connected Lie group whose semi-simple part has no compact factor. It is the aim of the present paper to give a description of the space $\mathcal{R}_1(\Gamma_0, G)$. We write G as a direct $G_1 \times G_2$ where G_1 is semi-simple and G_2 has no connected simple normal subgroup. Let R be the radical of G_2, S a semi-simple part of G_2, and N the maximal connected nilpotent normal subgroup of R. Denote by $\pi_i\colon G \to G_i$ $(i = 1, 2)$ the projections. Then $\pi_1(\Gamma_0)$ is a lattice in G_1 and so $\mathcal{R}_1(\pi_1(\Gamma_0), G_1)$ has a meaning. Our main results can be stated as follows:

(A) *Let A_0 denote the e-component of the group of automorphisms of the group $\Gamma_0 N$ with the compact-open topology, and $Ad_G S$ the group of all inner automorphisms of G induced by elements in S. Then, to each $r \in \mathcal{R}_1(\Gamma_0, G)$, there exist $r_1 \in \mathcal{R}_1(\pi_1(\Gamma_0), G_1)$, $\sigma \in Ad_G S$, $a \in A_0$ such that* $r(\gamma) = (r_1\pi_1(\gamma), \pi_2\sigma a(\gamma))$, $\gamma \in \Gamma_0$.

* Received July 17, 1962.

[1] Guggenheim Fellow and supported in part by National Science Foundation NSF-G-19079.

(B) *The map* $(r_1, \sigma, a) \to r$ *as defined in* (A) *is a covering map of* $\mathcal{R}_1(\pi_1(\Gamma_0), G_1) \times Ad_G S \times A_0 \to \mathcal{R}_1(\Gamma_0, G)$ *with finite covering group.*

Since G_1 is semi-simple without compact factor, the space $\mathcal{R}_1(\pi_1(\Gamma_0), G_1)$ is completely known [15], and so the above theorems give a description of the space $\mathcal{R}_1(\Gamma_0, G)$ in terms of the two groups of automorphisms $Ad_G S$ and A_0. As a Corollary, we have

(C) *If G is solvable, then the deformations of Γ_0 in G are in one-to-one correspondence with the elments in the e-component of the group of automorphisms of the non-connected Lie group $\Gamma_0 N$.*

2. Some properties of nilpotent Lie groups. Let N be a connected and simply-connected nilpotent Lie group. It possesses many particular properties of the real vector space when considered as a group. In this section, we shall give some such properties which will be used later.

Suppose that f is an automorphism of N and $\rho_f\colon N \to N$ is defined by $\rho_f(x) = f(x)x^{-1}$, $x \in N$. We shall call f unipotent if $\rho_f{}^m(N) = e$ for some integer $m > 0$ where e denotes the identity.

(2.1) *An automorphism f of N is unipotent if and only if the automorphism of the Lie algebra N of N induced by f is unipotent.*

Proof. When N is abelian, this is evident. For the general case, this can be verified directly by using induction on the length of the lower central series of N.

(2.2) *Let f be an automorphism of a nilpotent Lie algebra $\hat{N}$. If it is unipotent* mod $[\hat{N}, \hat{N}]$ (i.e., *there exists an integer m such that* $(f-1)^m\hat{N} \subset [\hat{N}, \hat{N}]$), *then it is unipotent.*

Proof. We shall prove it by induction. Suppose that (2.2) is valid for nilpotent Lie algebras of dimension less that that of $\hat{N}$. Let

$$\hat{N}^{(0)} = \hat{N}, \qquad \hat{N}^{(s)} = [\hat{N}, \hat{N}^{(s-1)}], \qquad\qquad s = 1, 2, 3, \cdots.$$

and t the largest integer for which $N^{(t)} \neq 0$. Denote by $\hat{N}^*$ the quotient $\hat{N}/\hat{N}^{(t)}$, $\pi\colon \hat{N} \to \hat{N}^*$ the projection, and f^* the automorphism of $\hat{N}^*$ induced by f. By assumption $(f-1)^m\hat{N} \subset [\hat{N}, \hat{N}]$ and so

$$(f^* - 1)^m \hat{N}^* \subset [\hat{N}^*, \hat{N}^*].$$

This tells us that f^* is unipotent mod $[\hat{N}^*, \hat{N}^*]$. It follows from the induction hypothesis that f^* is unipotent. There exists then an integer n such that $(f-1)^n\hat{N} \subset \hat{N}^{(t)}$.

Let $x, y \in \hat{N}$. Using the fact

$$(f-1)[x,y]=[(f-1)x,(f-1)y]+[(f-1)x,y]+[x,(f-1)y],$$

we can easily verify that, for any positive integer k, $(f-1)^k[x,y]$ is a linear combination of elements of the form

$$[(f-1)^q x,(f-1)^r y], \qquad 0\leqq q,r\leqq k;\, q+r\geqq k.$$

It follows that

$$(f-1)^{2n}[x,y]\in[\hat{N}^{(t)},\hat{N}]=0$$

and hence

$$(f-1)^{m+2n}(\hat{N})\subset(f-1)^{2n}[\hat{N},\hat{N}]=0.$$

The unipotency of f is thus proved.

(2.3) *Let Θ be a lattice in N and f an automorphism of N leaving Θ invariant. Suppose that $\rho_f{}^m(\Theta)=e$ for some integer m where ρ_f is defined by $\rho_f(x)=f(x)x^{-1}$. Then f is unipotent.*

Proof. Let $N^*=N/[N,N]$, $\pi\colon N\to N^*$ be the projection, $\Theta^*=\pi(\Theta)$, and f^* the automorphism of N^* induced by f. Then Θ^* is a lattice in N^*[9] and $\rho_{f^*}{}^m(\Theta^*)=e$. But N^* is isomorphic with a vector space and so f^* must be unipotent. In other words, f is unipotent mod $[N,N]$. Proposition (2.3) then follows directly from (2.1) and (2.2).

3. Some known results. For the sake of clarity, let us list in this section some relevant known results. Suppose G to be a topological group. A subgroup H of G will be said to have the Selberg property, or simply property (S), if for any neighborhood U of the identity e in G and every element g of G, there exists an integer $n>0$ with $g^n\in UHU$ [4]. As shown by Selberg [13], every lattice has the property (S).

(3.1) (A. Selberg; A. Borel). *Let G be a semi-simple Lie group without compact factor, and H a subgroup of G with the property (S). For any linear representation ρ of G, every invariant subspace of $\rho(H)$ is invariant under $\rho(G)$* [4].

As a consequence of this, we have

(3.2) *Let G, H have the same meaning as in* (3.1). *Any normal solvable subgroup L of H belongs to the center of G.*

Proof. Choose an irreducible representation ρ of G over a complex vector space V such that the kernel of ρ coincides with the center of G. Such a ρ always exists. If $\rho(L)=1$ then (3.2) is proved. If otherwise, let

$$\rho(L) = Q_0 \supset Q_1 \supset Q_2 \supset \cdots$$

be the derived series of $\rho(L)$ and let t be the largest integer for which $Q_t \neq 1$. Since Q_t is abelian, it has a non-trivial weight space V_ϕ, i. e. ϕ is a character of Q_t and

$$V_\phi = \{x \in V \colon g(x) = \phi(g)x, g \in Q_t\} \neq 0.$$

The number of such weight spaces is finite. Since $\rho(H)$ permutes these weight spaces, so does $\rho(G)$. But G is connected. Hence $\rho(G)(V_\phi) = V_\phi$, $V = V_\phi$, and Q_t contains only scalar matrices. From our choice of ρ, $Q_t = 1$. (3.2) is proved.

(3.3) (L. Auslander [2], H. Zassenhaus [16]). *Let G be a connected Lie group with simply-connected radical R, $\pi\colon G \to G/R$ the projection, and L a closed subgroup of G. If the e-component of L is solvable, then the e-component of the closure $\overline{\pi(L)}$ of $\pi(L)$ is also solvable.*

(3.4) (L. Auslander [2]). *Let G be a connected, simply-connected Lie group, R its radical, C the maximal connected normal compact subgroup of a semisimple part of G, and $\pi\colon G \to G/CR$ the projection. Suppose L to be a closed subgroup of G with its e-component solvable. If $\pi(L)$ has the property (S) in G/CR, then $\pi(L)$ is discrete.*

For application in Section 5, we need a generalized form (3.4) of a result of Auslander [2]: we do not assume L to be discrete. The proof is a refinement of the arguments in Auslander [2] and Zassenhaus [16]. However, since [2] has not appeared yet except as mimeographed notes, for completeness we give the proofs in full in the appendix rather than just the modifications necessary to our form of the proposition.

4. Lattices in the direct product of a semi-simple group and a vector space. As preparation for the general discussions, we shall consider lattices in a particular class of groups. Let Γ be a lattice in a direct product $G = G_1 \times G_2$ where G_1 is a connected semi-simple Lie group without compact factor and G_2 a vector space over the reals. As can be seen by examples, $\Gamma \cap G_1$ is not always a lattice in G_1. For latter use, we shall establish some properties of $\Gamma \cap G_1$.

Throughout this paper, for any real linear group L, we use $\mathfrak{A}(L)$ to denote the least real algebraic linear group which contains L.

(4.1) *Let $\Gamma_i = G_i \cap \Gamma$ and Γ_i' be the image of Γ under the projection $\pi_i\colon G \to G_i$ $(i = 1, 2)$. Then Γ_1', Γ_2 are lattices in G_1, G_2 respectively, and*

Γ_1 contains the commutator subgroup $[\Gamma_1', \Gamma_1']$. For any linear representation ρ of G_1, $\rho(G_1)$ is the e-component of $\mathfrak{A}(\rho(\Gamma_1))$.

Proof. The facts that Γ_1', Γ_2 are lattices in G_1, G_2 respectively follow directly from (3.4). Moreover since G_2 is abelian, we have $[\Gamma, \Gamma] \subset G_1$ and so

$$[\Gamma_1', \Gamma_1'] = \pi_1([\Gamma, \Gamma]) \subset \pi_1(\Gamma_1) = \Gamma_1.$$

Now let ρ be any representation of G_1. Then $\rho(\Gamma_1')$ is a subgroup of $\rho(G_1)$ having the property (S). Since $\rho(G_1)$ is semi-simple without compact factor, we have $\rho(G_1) \subset \mathfrak{A}(\rho(\Gamma_1')$ [4, p. 187]. It follows then

$$\rho(G_1) = [\rho(G_1), \rho(G_1)] \subset [\mathfrak{A}(\rho(\Gamma_1')), \mathfrak{A}(\rho(\Gamma_1'))] = \mathfrak{A}(\rho[\Gamma_1', \Gamma_1']) \subset \mathfrak{A}(\rho(\Gamma_1)).$$

But $\mathfrak{A}(\rho(\Gamma_1)) \subset \mathfrak{A}(\rho(G_1))$ and so $\mathfrak{A}(\rho(G_1)) = \mathfrak{A}(\rho(\Gamma_1))$. Therefore $\rho(G_1)$ must be the e-component of $\mathfrak{A}(\rho(\Gamma_1))$. (4.1) is thus proved.

(4.2) *Suppose that there exists a locally one-to-one representation $\rho\colon G \to GL(R, r)$ such that $\rho(\Gamma)$ contains only matrices with rational integers as entries. Then Γ_1 is a lattice in G_1, or what is the same, Γ_2' is discrete.*

Proof. Let $M = \rho(G)$, $M_1 = \rho(G_1)$ and let M_z, M_{1z} be respectively the sets of all integer matrices in M, M_1. Denote by Q the normalizer of M_{1z} in M_1. Since M_1 is a normal subgroup of M, Q must be invariant under $\operatorname{Ad} M_z$ and so QM_z forms a group. Let us first show that QM_z is discrete. From (4.1), $\mathfrak{A}(M_1) = \mathfrak{A}(\rho(\Gamma_1))$. But $\rho(\Gamma_1)$ contains only integer matrices. It follows that $\mathfrak{A}(\rho(\Gamma_1))$, and hence $\mathfrak{A}(M_1)$, is a real algebraic linear group defined over the field of rational numbers. Since M_1 is semi-simple, $\mathfrak{A}(M_1)/M_1$ is finite and therefore M_1/M_{1z} has a finite invariant measure [5]. Taking account of the results in [4], we see immediately that Q/M_{1z} is finite. Therefore QM_z/M_z is also finite. Since M_z is discrete, QM_z is also discrete.

Let γ be an element of Γ. It takes the form $\gamma = x_1 x_2$ where

$$x_1 = \pi_1(\gamma) \in G_1, \ x_2 = \pi_2(\gamma) \in G_2.$$

Since $\rho(\gamma)$ is an integer matrix, we have

$$\rho(x_1) M_{1z} \rho(x_1)^{-1} = \rho(\gamma) M_{1z} \rho(\gamma)^{-1} = M_{1z}.$$

This tells us that $\rho(\pi_1(\gamma)) \in Q$ for $\gamma \in \Gamma$, or what is the same $\rho(\Gamma_1') \subset Q$. It follows that

$$\rho(\Gamma_2') \subset \rho(\Gamma_1' \cdot \Gamma) \subset Q\rho(\Gamma) \subset QM_z$$

whence $\rho(\Gamma_2')$ is discrete. By assumption, ρ is locally one-to-one and so Γ_2' must be discrete. (4.2) is thus proved.

5. Projections of lattices. Let G be a connected and simply-connected Lie group, R its radical, and N the maximal connected nilpotent normal subgroup of R. Then G takes the form of a direct product $G = G_1 \times G_2$ where G_1 is semi-simple and G_2 has no connected semi-simple normal subgroup. On account of the simply-connectedness, G_2 is the semi-direct product SR where S is semi-simple. We note that S acts almost effectively on R under the adjoint action. Moreover [7, p. 109],

$$[G, R] \subset N, \qquad [G, G] \subset G_1 SN.$$

By using these inclusions, we can verify the following without difficulty:

(5.1) (a) $G/N \approx G_1 \times S \times R/N$. (b) *Let Z denote the center of N, and η the linear adjoint representation of G on the Lie algebra $\hat{N}$ of N. Then the mapping $S \to \eta(S)$ is a local isomorphism, and the e-component of the kernel of η is $G_1 \times Z$.*

After all these preliminaries, we are now in the position to prove a theorem

(5.2) THEOREM. *Suppose that G, G_1, G_2, S, R, N have the same meaning as above. Let Γ be a lattice in G and*

$$\phi\colon G \to G/R,\ \xi\colon G_2 \to G_2/R,\ \pi_i\colon G \to G_i \quad (i = 1, 2)$$

be the projections. If the semi-simple part of G has no compact factor, then (i) *the e-component of $\overline{\pi_2(\Gamma)}$ belongs to the center of N, and* (ii) *the image of Γ under the following projections are all discrete:*

$$G \to G/G_1R\ (\approx S),\ G \to G_1,\ G \to G/G_1SN\ (\approx R/N).$$

Proof. From (3.4), $\phi(\Gamma)$ is discrete and hence $R \cap \Gamma$ is a lattice in R. A theorem of Mostow [12] then asserts that $\circledast = N \cap (\Gamma \cap R) = \Gamma \cap N$ is a lattice in N whence the image of Γ under the projection $G \to G/N$ is discrete. With these in mind, let us discuss $\pi_2(\Gamma)$ which is, in general, non-discrete. Denote by η the linear adjoint representation Ad_N of G over the Lie algebra $\hat{N}$ of N. Then

$$\eta(\Gamma) = \eta(G_1\Gamma) = \eta(\pi_2(\Gamma)),\ \eta(\overline{\pi_2(\Gamma)}) \subset \overline{\eta(\Gamma)}.$$

The subgroup Θ is a lattice in N, and $\gamma\Theta\gamma^{-1} = \Theta$ for all $\gamma \in \Gamma$. From Malcev's results [9], there exists a base of $\hat{N}$ referred to which $\eta(\Gamma)$ contains only integer matrices, and hence $\eta(\Gamma)$ is discrete. Therefore $\eta(\overline{\pi_2(\Gamma)})$ must be discrete and the e-component of $\overline{\pi_2(\Gamma)}$ belongs to the kernel of η. From (5.1) it follows that the e-component of $\overline{\pi_2(\Gamma)}$ is contained in the center Z of N. (i) is thus proved. Applying (3.4) with $\overline{\pi_2(\Gamma)}$ taking the place of L, we see directly that $\xi(\overline{\pi(\Gamma)})$, and hence $\xi\pi_2(\Gamma)$, is discrete. In other words, the image of Γ under the projection $G \to G/G_1R$ $(= S)$ is discrete.

Let $\Delta = \phi(\Gamma)$ and let us identify G/R with $G_1 \times S$. We already know that Δ is a lattice in $G_1 \times S$ and its projection in S (which coincides with $\xi\pi_2(\Gamma)$) is discrete. Since both G_1 and S are semi-simple without compact factor, a result of Weil [15] asserts that the projection of Δ in the other factor G_1 must be also discrete. We note that this projection of Δ in G_1 is nothing but the image of Γ under the map $G \to G/SR$ $(= G_1)$.

Now it remains only to show the discreteness of the image of Γ under the projection $G \to G/G_1SN$ $(= R/N)$. For this purpose, let $N^* = N/[N,N]$. Since N is connected and simply-connected, N^* is a vector space over the reals. For any $g \in G$, $gNg^{-1} = N$, $g[N,N]g^{-1} = [N,N]$, and so g induces an automorphism $\eta^*(g)$ of N^*. We have thus a linear representation η^* of G over N^*. Let Θ^* be the image of $\Theta = N \cap \Gamma$ under the natural map $N \to N^*$. From Malcev's results [9], Θ^* is a lattice in N^*. Since $\eta^*(\gamma)\Theta^* = \Theta^*$, $\gamma \in \Gamma$, $\eta^*(\Gamma)$ is discrete. Let K be the kernel of η^*. Applying (5.1) to $G/[N,N]$, we can verify immediately that the e-component K_0 of K coincides with $G_1 \times N$. Therefore η^* can be written as the composite of the following two homomorphisms:

$$\psi\colon G \to G/K_0\ (\approx S \times R/N), \qquad f\colon G/K_0 \to \eta^*(G).$$

(We note here that $\eta^*(G)$ may have weaker topology than G/K, but is anyway a 1-1 homomorphic image of G/K). Evidently $\psi(\Gamma) \subset f^{-1}\eta^*(\Gamma)$. Since both $\eta^*(\Gamma)$ and the kernel of f are discrete, $\psi(\Gamma)$ must be also discrete, and hence a lattice in G/K_0.

Now we identify G/K_0 with $S \times R/N$. $\psi(\Gamma)$ is a lattice in this product. η^* induces a locally one-to-one representation ρ of $S \times R/N$ over the vector space N^*. Since $\rho\psi(\Gamma) = \eta^*(\Gamma)$ contains only integer matrices, we know from (4.2) that the image of $\psi(\Gamma)$ under the projection $S \times R/N \to R/N$ is discrete. In other words, the image of Γ under the projection $G \to G/G_1SN$ is discrete. This completes the proof.

As an immediate consequence of (5.2), we have

(5.3) *Let* $\lambda\colon G\to S$, $\pi_1\colon G\to G_1$, $\pi\colon G\to R/N$ *be the natural projections. Then, in each of the following pairs of groups, the former is a lattice in the latter:*

$$(\lambda(\Gamma),S),\ (\pi_1(\Gamma),G_1),\ (\pi(\Gamma),R/N),\ (G_1R\cap\Gamma,G_1R),\ (G_2\cap\Gamma,G_2),\ (G_1SN\cap\Gamma,G_1SN).$$

6. Some density properties of Γ. Suppose that G, G_1, S, R, N, $\Gamma,\cdots$ have the same meaning as in Section 5, and assume that the semi-simple part of G has no compact factor. Let us recall that, for any real linear group, $\mathcal{A}(L)$ denotes the least real algebraic linear group containing L.

(6.1) *Let ρ be a real linear representative of G such that $\rho(N)$ contains only unipotent matrices. Then $\rho(G_1\times SN)\subset\mathcal{A}(\rho(\Gamma))$.*

Proof. Since $N\cap\Gamma$ is a lattice in N, there exists a base $X_1,\cdots,X_s$ of the Lie algebra $\hat{N}$ of N such that

$$\gamma_i=\exp X_i\in N\cap\Gamma \qquad (i=1,2,\cdots,s).$$

By assumption, $\rho(\exp tX_i)$ is unipotent for any real t and so $\mathcal{A}(\rho(\gamma_i))=\{\rho(\exp tX_i):t=\text{real}\}$. It follows that

$$\rho(N)\subset\mathcal{A}(\rho(\Gamma)). \tag{6.2}$$

Let $\lambda\colon G\to G/N$ be the projection and let us identify G/N and the product $(G_1\times S)\times R/N$. From Theorem (5.2), $\lambda(\Gamma)$ is a lattice in $(G_1\times S)\times R/N$. Since $(G_1\times S)$ is semi-simple without compact factor and R/N a vector group, we can apply (4.1), obtaining

$$\rho(G_1\times S)\subset\mathcal{A}(\rho(\lambda(\Gamma)\cap(G_1\times S)).$$

Up to our identification, λ is given by

$$\lambda(g_1)=g_1,\ \lambda(s)=s,\ \lambda(r)=rN$$
$$g_1\in G_1,\ s\in S,\ r\in R.$$

It follows that $\lambda(\Gamma)\cap(G_1\times S)=\Gamma N\cap(G_1\times S)$ and therefore

$$\rho(G_1\times S)\subset\mathcal{A}(\rho(\Gamma N)).$$

Combining this with (6.2), we have

$$\rho(G_1SN)\subset\mathcal{A}(\rho(\Gamma N))=\mathcal{A}(\rho(\Gamma))$$

and (6.1) is proved.

(6.3) *Let α be an automorphism of G leaving Γ pointwise fixed. Then α leaves $G_1 \times SN$ pointwise fixed.*

Proof. Let $\dot{\alpha}$ denote the automorphism of the Lie algebra $\hat{G}$ of G induced by α. Since $\dot{\alpha} \cdot \mathrm{Ad}\,\gamma = \mathrm{Ad}\,\gamma \cdot \alpha$ for all $\gamma \in \Gamma$, it follows from (6.1) that $\dot{\alpha} \cdot \mathrm{Ad}\,g = \mathrm{Ad}\,g \cdot \dot{\alpha}$ for all $g \in G_1 \times SN$, whence $g^{-1}\alpha(g)$ belongs to the center C of G. We can easily verify that the map $q\colon G_1 \times S \to C$ defined by $q(g) = g^{-1}\alpha(g)$, $g \in G_1 \times S$, is a homomorphism. Since $G_1 \times S$ is semi-simple and C abelian, q must be trivial, i.e., $\alpha(g) = g$, $g \in G_1 \times S$. On the other hand, $\Gamma \cap N$ is a lattice in N and α leaves $\Gamma \cap N$ pointwise fixed. It follows that $\alpha(n) = n$ for all $n \in N$. (6.3) is proved.

(6.4) COROLLARY. *The e-component of $\overline{\pi_2(\Gamma)}$ belongs to the centralizer of SN in G_2.*

Proof. The intersection $G_2 \cap \Gamma$ is a normal subgroup of Γ. It follows that $G_2 \cap \Gamma$ is a normal subgroup of $\pi_2(\Gamma)$, and hence, of $\overline{\pi_2(\Gamma)}$. Therefore the e-component of $\overline{\pi_2(\Gamma)}$ must centralize $G_2 \cap \Gamma$. But by Theorem (5.2), $\pi_1(\Gamma)$ is discrete and so $G_2 \cap \Gamma$ is a lattice in G_2. Applying (6.3) with G_2 taking the place of G, we know that the e-component of $\overline{\pi_2(\Gamma)}$ must centralize SN. (6.4) is thus proved.

7. Stability of subgroups under deformation. Let Γ_0 be a fixed lattice in a Lie group G. Following Weil, we use $\mathcal{R}(\Gamma_0, G)$ to denote the space, with the compact-open topology, of all the isomorphisms $r\colon \Gamma_0 \to G$ such that $r(\Gamma_0)$ is a lattice in G. The connected component of $\mathcal{R}(\Gamma_0, G)$ which contains the identity map $r_0\colon \Gamma_0 \to G$ will be denoted by $\mathcal{R}_1(\Gamma_0, G)$. In fact, Γ_0 does not play any special role in $\mathcal{R}_1(\Gamma_0, G)$. If we take $r \in \mathcal{R}_1(\Gamma_0, G)$ and put $\Gamma = r(\Gamma_0)$, then $\mathcal{R}_1(\Gamma, G) = \mathcal{R}_1(\Gamma_0, G)$.

Definition. Let L be a closed subgroup of G, and Γ_0 a lattice in G. L is called stable under deformation of Γ_0 if the subgroup $r^{-1}(L)$ of Γ_0 is independent of the choice of r in $\mathcal{R}_1(\Gamma_0, G)$. This condition is the same as $r(\Gamma_0 \cap L) = L \cap r(\Gamma_0)$, $r \in \mathcal{R}_1(\Gamma_0, G)$.

If L is normal and $\pi\colon G \to G/L$ denotes the projection, then we can see immediately that L is stable if and only if the kernel of $\pi r\colon \Gamma_0 \to G/L$ does not depend on the choice of $r \in \mathcal{R}_1(\Gamma_0, G)$.

(7.1) *If, for each $r \in \mathcal{R}_1(\Gamma_0, G)$,* there is a neighborhood $\mathcal{V}$ *of r in $\mathcal{R}_1(\Gamma_0, G)$ such that $r^{-1}(L) \subset s^{-1}(L)$ for all $s \in \mathcal{V}$, then L is stable under deformation of Γ_0.*

Proof. Let r_1 be an arbitrary but fixed element in $\mathfrak{R}_1(\Gamma_0, G)$. From the assumption, we see easily that the set

$$\mathfrak{W} = \{r : r \in \mathfrak{R}_1(\Gamma_0, G), r_1^{-1}(L) \subset r^{-1}(L)\}$$

is an open subset of $\mathfrak{R}_1(\Gamma_0, G)$. Since the condition $r_1^{-1}(L) \subset r^{-1}(L)$ is the same as $rr_1^{-1}(L) \subset L$, $\mathfrak{W}$ must be closed, and hence $\mathfrak{W} = \mathfrak{R}_1(\Gamma_0, G)$. This tells us that $r_1^{-1}(L) \subset r^{-1}(L)$ for all $r \in \mathfrak{R}_1(\Gamma_0, G)$. But r_1 is arbitrarily chosen, and so $r^{-1}(L) = r'^{-1}(L)$ for all $r, r' \in \mathfrak{R}_1(\Gamma_0, G)$. (7.1) is thus proved.

(7.2) *Let L be a closed normal subgroup of G with G/L semi-simple and having no compact factor. If $\pi r(\Gamma_0)$ are discrete for all $r \in \mathfrak{R}_1(\Gamma_0, G)$ where $\pi: G \to G/L$ denotes the projection, then L is stable under the deformation of Γ_0.*

Proof. We find it convenient first to make two conventions. Let Q be a subset of G, and m a positive integer. The totality of commutators of the form

$$[q_1[q_2[\cdots[q_{m-1}, q_m]\cdots]]], \quad q_i \in Q$$

will be denoted by $\mathfrak{L}_m(Q)$. By $\lim_{m\to\infty} \mathfrak{L}(Q) = e$, we mean that given any neighborhood U of e in G, there exists an integer n such that $\mathfrak{L}_m(Q) \subset U$ for $m \geqq n$.

Now let us prove (7.2). From (7.1), it suffices to show that each r has a neighborhood $\mathfrak{V}$ in $\mathfrak{R}_1(\Gamma_0, G)$ such that $r^{-1}(L) \subset s^{-1}(L)$ for al s $\in \mathfrak{V}$. For simplicity of notation and without loss of generality, we can take r to be r_0. By assumption, $\Theta_0 = \Gamma_0 \cap L$ is a lattice in L. Hence it has a finite set of generators $\beta_1, \beta_2, \cdots, \beta_a$. Choose a neighborhood U_1 of e in G/L such that $\lim \mathfrak{L}_m(U_1) = e$, and U_1 contains no central element of G/L except e. Such a U_1 always exists [16]. Let $U = \pi^{-1}(U_1)$ where $\pi: G \to G/L$ denotes the projection, and

$$\mathfrak{V} = \{r \in \mathfrak{R}_1(\Gamma_0, G) : \beta_i^{-1}r(\beta_i) \in U, i = 1, 2, \cdots, a\}.$$

This $\mathfrak{V}$ is a neighborhood of r_0 in $\mathfrak{R}_1(\Gamma_0, G)$. Let r be an arbitrary element in $\mathfrak{V}$ and $F = \{\beta_1, \beta_2, \cdots, \beta_a\}$. Then $\pi(r(\beta_i)) \in \pi(\beta_i U) = \pi(U) = U_1$ and hence

$$\lim_{m\to\infty} \mathfrak{L}_m(\pi r(F)) \subset \lim_{m\to\infty} \mathfrak{L}_m(U_1) = e.$$

But $\pi r(\Theta_0)$ is discrete and $\mathfrak{L}_m(\pi r(F)) \subset \pi r(\Theta_0)$. There exists then an integer n such that

$$\mathcal{L}_n(\pi r(F)) = e.$$

It follows [16] that the subgroup $\pi r(\Theta_0)$ which is generated by $\pi r(F)$ must be nilpotent. Now we have the following situation: G/L is a semi-simple Lie group without compact factor; $\pi r(\Gamma_0)$ is a lattice in G/L and $\pi r(\Theta_0)$ is a nilpotent normal subgroup of $\pi r(\Gamma_0)$. From (3.2), $\pi r(\Theta_0)$ belongs to the center of G/L. But U_1 does not contain any central element of G/L except the identity. It follows that $\pi r(\beta_i) = e$ and then $\pi r(\Theta_0) = e$. In other words, $r(\Theta_0) \subset L$. Thus we have shown the existence of the neighborhood $\mathfrak{V}$ of r_0 such that $r^{-1}(L) \supset \Theta_0 = r_0^{-1}(L)$ for all $r \in \mathfrak{V}$. This proves our proposition (7.2).

(7.3) *Suppose that G, G_1, G_2, S, R, N have the same meaning as in Section 5, Γ_0 is a lattice in G, and the semi-simple part of G has no compact factor. Then the subgroups N, R, SR, G_1R are stable under deformation of Γ_0.*

Proof. The fact that SR, G_1R are stable follows immediately from Theorem (5.2) and Proposition (7.2). Since $R = SR \cap G_1R$, R is also stable.

On account of (7.1), to prove the stability of N, it suffices to find, for each $r \in \mathfrak{R}_1(\Gamma_0, G)$, a neighborhood $\mathfrak{V}$ of r in $\mathfrak{R}_1(\Gamma_0, G)$ such that

$$r^{-1}(N) \subset s^{-1}(N), \quad s \in \mathfrak{V}.$$

Since r_0 does not play any particular role in the set $\mathfrak{R}_1(\Gamma_0, G)$, it will be sufficient to construct the neighborhood $\mathfrak{V}$ for r_0.

Choose a neighborhood U of e in G with the following properties: (i) to each $x \in U$, there corresponds a unique element X of the Lie algebra $\hat{G}$ of G such that $\exp X = x$, $\exp tX \in U$, $0 \leqq t \leqq 1$; (ii) if $x \in R$, then X belongs to the Lie algebra $\hat{R}$ of R; (iii) if $\operatorname{Ad} x$ is unipotent on the Lie algebra $\hat{N}$ of N, then $\operatorname{Ad} X$ is nilpotent on $\hat{N}$. Let $\beta_1, \beta_2, \cdots, \beta_a$ be a set of generators of $\Theta_0 = N \cap \Gamma_0$, and

$$\mathfrak{V} = \{r \in \mathfrak{R}_1(\Gamma_0, G) : \beta_i^{-1} r(\beta_i) \in U, i = 1, 2, \cdots, a\}.$$

This $\mathfrak{V}$ is a neighborhood of r_0 in $\mathfrak{R}_1(\Gamma_0, G)$. Take $r \in \mathfrak{V}$ and put $\Theta = r^{-1}(N)$. Since N is a nilpotent normal subgroup of G, both Θ and Θ_0 are nilpotent normal subgroups of Γ_0. There exists then an integer m such that

$$\underbrace{[\beta_i, [\beta_i, [\cdots [\beta_i,}_{m} \Theta] \cdots]]] = e$$

whence

$$\underbrace{[r(\beta_i), [r(\beta_i), [\cdots [r(\beta_i),}_{m} r(\Theta)] \cdots]]] = e.$$

By (5.2), $r(\mathfrak{G}) = N \cap r(\Gamma_0)$ is a lattice in N, and so by (2.3)

$$\underbrace{[r(\beta_i), [r(\beta_i), [\cdots [r(\beta_i)}_{m}, N] \cdots]]] = e$$

It follows then that, for each i, the subgroup M_i generated by $r(\beta_i)$ and N is nilpotent. Since $\beta_i^{-1}r(\beta_i) \in M_i$, the automorphism of N given by $x \to \beta_i^{-1}r(\beta_i)xr(\beta_i)^{-1}\beta_i$, $x \in N$ is unipotent in the sense of Section 2. (2.1) then tells us that $\mathrm{Ad}\,\beta_i^{-1}r(\beta_i)$ is unipotent on $\hat{N}$.

By our choice of r, $\beta_i^{-1}r(\beta_i) \in U$. Let Y be the unique element of $\hat{G}$ with $\exp Y = \beta_i^{-1}r(\beta_i)$ as given in the definition of U. Because of the property (iii), $\mathrm{Ad}\,Y$ is nilpotent on $\hat{N}$. On the other hand, we already know that R is stable and so

$$\beta_i^{-1}r(\beta_i) \in N \cdot r(N \cap \Gamma_0) \subset N \cdot r(R \cap \Gamma_0) \subset NR = R.$$

From property (ii), Y must belong to the Lie algebra $\hat{R}$ of R. But $\mathrm{Ad}\,Y$ is nilpotent on $\hat{N}$ and so $Y \in \hat{N}$, $\beta_i^{-1}r(\beta_i) \in N$. It follows that $r(\beta_i) \in N$, $r(\mathfrak{G}_0) \subset N$, or what is the same,

$$r_0^{-1}(N) = \mathfrak{G}_0 \subset r^{-1}(N); \quad r \in \mathfrak{V}.$$

This proves that N is stable under deformation of Γ_0. The proof is thus completed.

As an immediate consequence of (5.2) and (7.3), we have

(7.4) *Let*

$$\pi\colon G \to R/N\ (= G/G_1SN); \quad \lambda\colon G \to S\ (= G/G_1R), \quad \pi_1\colon G \to G_1$$

be the projections. Then $\pi r(\Gamma_0 \cap G_1SN)$, $\lambda r(\Gamma_0 \cap G_1R)$ *and* $\pi_1 r(\Gamma_0 \cap G_2)$ *are trivial for all* $r \in \mathfrak{R}_1(\Gamma_0, G)$.

8. The projection $\pi\colon G \to R/N$. It is the aim of this section to show that $\pi(\gamma) = \pi r(\gamma)$, $\gamma \in \Gamma_0$, $r \in \mathfrak{R}_1(\Gamma_0, G)$ where $\pi\colon G \to R/N$ denotes the projection. We shall first prove this for solvable G and then extend it to the general case.

(8.1) (a) *For any semi-simple element* $A \in GL(R, n)$, *there exists a neighborhood* U *of the identity in* $GL(R, n)$ *such that if* $B \in U$ *and if* BAB^{-1} *commutes with* A, *then* $AB = BA$. (b) *If* $A, B \in GL(R, n)$ *have the same semi-simple part and commute with each other; then* $A^{-1}B$ *is unipotent.*

Proof. Part (a) has been proved by Frobenius [8]. To see Part (b), let A_u, B_u be the unipotent parts of A, B respectively. Then $A^{-1}B = A_u^{-1}B_u$.

But A_u^{-1}, B_u are unipotent and commute with each other, and so $A_u^{-1}B_u$ is unipotent.

(8.2) *Let R be a connected, simply-connected solvable Lie group, N the maximal connected nilpotent normal subgroup and $\pi: R \to R/N$ the projection. For any lattice Φ_0 in R, $\pi(\gamma) = \pi(r(\gamma))$, $\gamma \in \Phi_0$, $r \in \mathcal{R}_1(\Phi_0, R)$.*

Proof. Since the set $\{r \in \mathcal{R}_1(\Phi_0, R) : \pi(\gamma) = \pi(r(\gamma)), \gamma \in \Phi_0\}$ is closed in $\mathcal{R}_1(\Phi_0, R)$, it suffices to show that the identity map $r_0: \Phi_0 \to R$ has a neighborhood $\mathcal{V}$ such that $\pi(\gamma) = \pi r(\gamma)$, $\gamma \in \Phi_0$, $r \in \mathcal{V}$.

Let $\gamma_1, \gamma_2, \cdots, \gamma_a$ be a set of generators of Φ_0. Choose a neighborhood W of e in R such that (i) to each $x \in W$, there corresponds a unique element X in the Lie algebra $\hat{R}$ of R such that $\exp X = x$, $\exp tX \in W$, $0 \leqq t \leqq 1$, and (ii) if $\operatorname{Ad} x$ is unipotent, then $\operatorname{Ad} X$ is nilpotent. Put

$$\mathcal{V}_1 = \{r : \gamma_i^{-1} r(\gamma_i) \in W, i = 1, 2, \cdots, a\}.$$

This is a neighborhood of r_0.

Now let us find another neighborhood of r_0. We note that, for $g \in R$, $\operatorname{Ad} g$ leaves invariant both N and $[N, N]$. Therefore $\operatorname{Ad} g$ induces an automorphism $\eta^*(g)$ of the linear space $N^* = N/[N, N]$. The kernel of this representation has N as its e-component, and hence $\eta^*(R)$ is abelian. Let Θ_0 be the intersection $N \cap \Phi_0$. From (5.2), Θ_0 is a lattice in N. Since N is stable under deformation of Φ_0, $r(\Theta_0) = N \cap r(\Phi_0)$ are all lattices in N, $r \in \mathcal{R}_1(\Phi_0, R)$. By Malcev's results [9], to each r there exists a unique automorphism t_r of N such that $r(\beta) = t_r(\beta)$, $\beta \in \Theta_0$. Let t_r^* denote the automorphism of N^* induced by t_r. The map $r \to t_r^*$ is evidently continuous. Since $t_{r_0}^*$ is the identity, from (8.1a) we can find a neighborhood $\mathcal{V}_2$ of r_0 in $\mathcal{R}_1(\Phi_0, R)$ such that if $r \in \mathcal{V}_2$ and if $t_r^* \eta^*(\gamma_i)_s t_r^{*-1}$ commutes with $\eta^*(\gamma_i)_s$ $(i = 1, 2, \cdots, a)$, then $t_r^* \eta^*(\gamma_i)_s = \eta^*(\gamma_i)_s t_r^*$ where $\eta^*(\gamma_i)_s$ denotes the semi-simple part of $\eta^*(\gamma_i)$.

Now let us show that the neighborhood $\mathcal{V}_1 \cap \mathcal{V}_2$ has the required property. Suppose $r \in \mathcal{V}_1 \cap \mathcal{V}_2$ and $\theta: N \to N^*$ to be the projection. For $\beta \in \Theta_0$, we have

$$t_r(\gamma_i \beta \gamma_i^{-1}) = r(\gamma_i) r(\beta) r(\gamma_i)^{-1} = r(\gamma_i) t(\beta) r(\gamma_i)^{-1}$$

whence

$$t_r^* \eta^*(\gamma_i) \theta(\beta) = \eta^*(r(\gamma_i)) t_r^* \theta(\beta).$$

The image $\theta(\Theta_0)$ is a lattice in N^*, and therefore $t_r^* \eta^*(\gamma_i) = \eta^*(r(\gamma_i)) t_r^*$. It follows that

$$t_r^* \eta^*(\gamma_i)_s t_r^{*-1} = \eta^*(r(\gamma_i))_s$$

where the lower index s denotes the semi-simple part. Since $\eta^*(R)$ is abelian, $\eta^*(\gamma_i)_s$ commutes with $\eta^*(r(\gamma_i))_s$. From our choice of $\mathfrak{V}_2$, we have then $t_r^*\eta^*(\gamma_i)_s = \eta^*(\gamma_i)_s t_r^*$ whence $\eta^*(\gamma_i)_s = \eta^*(r(\gamma_i))_s$. By (8.1b), the product

$$\eta^*(\gamma_i)^{-1}\eta^*(r(\gamma_i)) = \eta^*(\gamma_i^{-1}r(\gamma_i))$$

is unipotent. (2.2) then implies that $\mathrm{Ad}(\gamma_i^{-1}r(\gamma_i))$ is unipotent on the Lie algebra $\hat{N}$ of N. Let X be the element in the Lie algebra of R with $\exp X = \gamma_i^{-1}r(\gamma_i)$ as given in the definition of $\mathfrak{V}_1$. Then $\mathrm{ad}\, X$ is nilpotent on $\hat{N}$ and hence $X \in \hat{N}$. It follows that $r(\gamma_i) \in \gamma_i N$ and $\pi(\gamma_i) = \pi r(\gamma_i)$. Since $\gamma_1, \cdots, \gamma_a$ generates Φ_0, $\pi(\gamma) = \pi r(\gamma)$ for all γ in Φ_0. (8.2) is thus proved.

(8.3) *Suppose that G, S, R, N, Γ_0, $\pi, \cdots$ have the same meaning as in Section 7. Then $\pi(\gamma) = \pi r(\gamma)$, $\gamma \in \Gamma_0$, $r \in \mathcal{R}_1(\Gamma_0, G)$.*

Proof. The quotient R/N, being connected, simply-connected and abelian, has an intrinsic real linear structure. Let us take R/N to be a real vector space. By (5.2), $\pi(\Gamma_0 \cap R)$ is a lattice in R/N. We can find $\gamma_1, \gamma_2, \cdots, \gamma_m$ in $\Gamma_0 \cap R$ such that $\pi(\gamma_1), \cdots, \pi(\gamma_m)$ form a base of R/N. From (8.2), $\pi r(\gamma_i) = \pi(\gamma_i)$, $i = 1, 2, \cdots, m$. Let Q be the totality of linear combinations of $\pi(\gamma_1), \cdots, \pi(\gamma_m)$ with rational coefficients. Since $\pi(\gamma_i) \in \pi r(\Gamma_0)$ and $\pi r(\Gamma_0)$ is discrete, we have $\pi r(\Gamma_0) \subset Q$. Now fix an element $\gamma \in \Gamma_0$. The set $\{\pi r(\gamma) : r \in \mathcal{R}_1(\Gamma_0, G)\}$ is connected; but Q is totally disconnected. Therefore $\pi r(\gamma) = \pi(\gamma)$ for all $r \in \mathcal{R}_1(\Gamma_0, G)$. (8.3) is thus proved.

9. The projection $\lambda: G \to S$. From (5.2), the image $\lambda(\Gamma_0)$ under the projection λ: $G \to S$ $(= G/G_1R)$ is a lattice in S. Moreover, G_1R is stable under deformation of Γ_0, and so each $r \in \mathcal{R}_1(\Gamma_0, G)$ induces a deformation of $\lambda(\Gamma_0)$ in S. In this section, we shall show that these induced deformations of $\lambda(\Gamma_0)$ can all be obtained by inner automorphisms of S.

Let us recall the representation η^* which we have used in Section 5. Since both N and $[N, N]$ are normal subgroups of G, each $g \in G$ induces an automorphism $\eta^*(g)$ of the quotient $N^* = N/[N, N]$ defined by $x[N, N] \to gxg^{-1}[N, N]$, $x \in N$. N^* is a vector space and so η^* is a linear representation. The kernel of η^* contains $G_1 \times N$ as its e-component. Therefore $S \to \eta^*(S)$ is a covering map.

Let $\Theta_0 = \Gamma_0 \cap N$ and $r \in \mathcal{R}_1(\Gamma_0, G)$. From (5.2) and (7.3), $r(\Theta_0)$ belongs to N and is a lattice in N. We know then from Malcev's results [9] that there exists a unique automorphism t_r of N such that $t_r(\beta) = r(\beta)$, $\beta \in \Theta_0$. It follows then

$$t_r(\gamma\beta\gamma^{-1}) = r(\gamma\beta\gamma^{-1}) = r(\gamma)t_r(\beta)r(\gamma)^{-1}, \quad \gamma \in \Gamma_0, \quad \beta \in \Theta_0.$$

Let t_r^* be the automorphism of N^* induced by t_r, and $\theta: N \to N^*$ the projection. Then

$$t_r^*\eta^*(\gamma)\theta(\beta) = \theta(r(\gamma)t_r(\beta)r(\gamma)^{-1}) = \eta^*(r(\gamma))t_r^*\theta(\beta).$$

This means that the two automorphisms $t_r^*\eta^*(\gamma)$ and $\eta^*(r(\gamma))t_r^*$ of N^* are identical on $\theta(\circledS_0)$. But $\theta(\circledS_0)$ is a lattice in N^*, and so they must be identical on the entire N^*, i.e.,

$$(9.1) \qquad \eta^*(r(\gamma)) = t_r^*\eta^*(\gamma)t_r^{*-1}.$$

From (5.2), the intersection $\Psi_0 = \Gamma_0 \cap G_1SN$ is a lattice in G_1SN, and from (8.3), $r(\Psi_0) \subset G_1SN$, $r \in \mathcal{R}_1(\Gamma_0, G)$. Bearing these in mind and applying (6.1) with G_1SN, Ψ_0, η^* taking the places of G, Γ, ρ, we obtain

$$\begin{aligned} t_r^*\eta^*(S)t_r^{*-1} \subset t_r^*\mathcal{A}(\eta^*(\Psi_0))t_r^{*-1} &= \mathcal{A}(t_r^*\eta^*(\Psi_0)t_r^{*-1}) \\ &= \mathcal{A}(\eta^*(r(\Psi_0))) \subset \mathcal{A}(\eta^*(G_1SN)) = \mathcal{A}(\eta^*(S)). \end{aligned}$$

Since S is semi-simple, $\eta^*(S)$ is the e-component of $\mathcal{A}(\eta^*(S))$. Hence

$$t_r^*\eta^*(S)t_r^{*-1} = \eta^*(S)$$

and therefore t_r^* induces an automorphism σ_r^* of $\eta^*(S)$. The map $r \to \sigma_r^*$ is continuous. It follows then from the connectedness of $\mathcal{R}_1(\Gamma_0, G)$ that σ_r^* is an inner automorphism of $\eta^*(S)$.

We know that S is semi-simple, and the map $\eta^*|_S: S \to \eta^*(S)$ is a local isomorphism. There exists then a natural one-to-one correspondence between inner automorphisms of S and those of $\eta^*(S)$. Denote by σ_r the automorphism of S corresponding to σ_r^*. Then $\eta^*\sigma_r(s) = \sigma_r^*\eta^*(s)$ for all $s \in S$, and then

$$(9.2) \qquad \eta^*\sigma_r\lambda = \sigma_r^*\eta^*\lambda.$$

Moreover, we can easily see that

$$(9.3) \qquad \eta^*(g) = \eta^*\lambda(g), \quad g \in G_1SN.$$

Now take an arbitrary but fixed element β in the intersection $\Psi_0 = \Gamma_0 \cap G_1SN$. Since G_1SN is stable under deformation of Γ_0, $r(\beta) \in G_1SN$, $r \in \mathcal{R}_1(\Gamma_0, G)$. From (9.1), (9.2), (9.3) and the definition of σ_r^*, we have

$$\eta^*\lambda r(\beta) = \eta^*r(\beta) = t_r^*\eta^*(\beta)t_r^{*-1} = \sigma_r^*\eta^*(\beta) = \sigma_r^*\eta^*\lambda(\beta) = \eta^*\sigma_r\lambda(\beta).$$

Let Z^* be the kernel of the local isomorphism $\eta^*|_S: S \to \eta^*(S)$. Then the above equality tells us that $\lambda r(\beta)$ and $\sigma_r\lambda(\beta)$ can only differ by an element $z(r)$ in Z^*. In other words,

$$\sigma_r\lambda(\beta)=\lambda(r(\beta))z(r), \quad z(r)\in Z^*.$$

As r runs through $\mathcal{R}_1(\Gamma_0,G)$, $z(r)$ describes a connected set. But Z^* is discrete, and so $z(r)=z(r_0)=e$. Hence

$$(9.4) \qquad \sigma_r\lambda(\beta)=\lambda r(\beta),\ r\in\mathcal{R}_1(\Gamma_0,G),\ \beta\in\Psi_0.$$

Now let us show that formula (9.4) holds not only for β in Ψ_0 but also for any element in Γ_0. Suppose γ to be arbitrary but fixed element in Γ_0. Put $s_1=\sigma_r\lambda(\gamma)$, $s_2=\lambda r(\beta)$, $s_3=s_2^{-1}s_1$. Since Ψ_0 is a normal subgroup of Γ_0, $\gamma\beta\gamma^{-1}\in\Psi_0$ for all $\beta\in\Psi_0$. From (9.4), we have then

$$s_2(\lambda r(\beta))s_2^{-1}=\lambda r(\gamma\beta\gamma^{-1})=\sigma_r\lambda(\gamma\beta\gamma^{-1})=s_1(\sigma_r\lambda(\beta))s_1^{-1}=s_1(\lambda r(\beta))s_1^{-1},$$

or what is the same,

$$(9.5) \qquad s_3(\lambda r(\beta))s_3^{-1}=\lambda r(\beta),\ \beta\in\Psi_0,\ r\in\mathcal{R}_1(\Gamma_0,G).$$

On account of (5.3) and (8.3), $r(\Psi_0)$ is a lattice in G_1SN, and so $\lambda r(\Psi_0)$ is a lattice in S. Now S is a semi-simple Lie group without compact factor, and s_3 is an element which centralizes the lattice $\lambda r(\Psi_0)$ in S. It follows then that s_3 belongs to the center of S. Since s_3 depends continuously on r. s_3 describes a connected set when r runs through $\mathcal{R}_1(\Gamma_0,G)$. Therefore $s_3=e$, whence

$$\sigma_r\lambda(\gamma)=s_1=s_2=\lambda r(\gamma),\ \gamma\in\Gamma_0,\ r\in\mathcal{R}_1(\Gamma_0,G).$$

Summarizing the above results, we have

(9.6) *To each $r\in\mathcal{R}_1(\Gamma_0,G)$, there corresponds an inner automorphism σ_r of S such that $\sigma_r\lambda(\gamma)=\lambda r(\gamma)$, for all $\gamma\in\Gamma_0$. Moreover, the map $r\to\sigma_r$ is continuous.*

10. A subspace of $\mathcal{R}_1(\Gamma_0,G)$. The main purpose of this paper is to give a description of deformations or Γ_0 in G. In this section, we shall show that our problem can be reduced to the description of a subspace of $\mathcal{R}_1(\Gamma_0,G)$. Let us first state the following rather evident proposition.

(10.1) *Let M be a subgroup of a direct product $L=L_1\times L_2$, and $\phi\colon L\to L_1$ the projection. If $\phi(M)$ and $M\cap L_2$ are lattices in L_1 and L_2 respectively, then M is a lattice in L.*

Now we assume that G, G_1, G_2, S, R, N, $\Gamma_0,\cdots$ have the same meaning as before and that the semi-simple part of G has no compact factor. Put

$$\mathcal{R}_2=\{r\in\mathcal{R}_1(\Gamma_0,G):\gamma^{-1}r(\gamma)\in G_2,\gamma\in\Gamma_0\},$$

and $\Gamma_1' = \pi_1(\Gamma_0)$. From (5.3), we know that Γ_1' is a lattice in G_1, and hence $\mathfrak{R}_1(\Gamma_1', G_1)$ has a meaning. Suppose that $q \in \mathfrak{R}_1(\Gamma_1', G)$ and $u \in \mathfrak{R}_2$. Let us consider the map $r \colon \Gamma_0 \to G$ defined by

$$r(\gamma) = (q\pi_1(\gamma), \pi_2 u(\gamma)), \gamma \in \Gamma_0. \tag{10.2}$$

By using (10.1) and the definition of $\mathfrak{R}_2$, we can easily verify that $r(\Gamma_0)$ is a lattice in G. To see that $r \in \mathfrak{R}_1(\Gamma_0, G)$, let us take an arc q_t $(0 \leqq t \leqq 1)$ in $\mathfrak{R}_1(\Gamma_1', G)$ with $q_1 = q$ and q_0 being the identical map $\Gamma_1' \to G_1$. Define $V_t \colon \Gamma_0 \to G$ by

$$V_t(\gamma) = (q_t\pi_1(\gamma), \pi_2 u(\gamma)), \quad 0 \leqq t \leqq 1.$$

Then V_t $(0 \leqq t \leqq 1)$ is an arc in $\mathfrak{R}(\Gamma_0, G)$ with $V_1 = r$, $V_0 = u$ where $\mathfrak{R}(\Gamma_0, G)$ denotes the space of all isomorphisms $\omega \colon \Gamma_0 \to G$ such that $\omega(\Gamma_0)$ is a lattice in G. Since u is an element of $\mathfrak{R}_1(\Gamma_0, G)$ so must be r. Hence to each element q of $\mathfrak{R}_1(\Gamma_1', G_1)$ and each element u of $\mathfrak{R}_2$, there corresponds an element r of $\mathfrak{R}_1(\Gamma_0, G)$ as given by (10.2). Let us denote this correspondence by $h \colon \mathfrak{R}_1(\Gamma_1', G_1) \times \mathfrak{R}_2 \to \mathfrak{R}_1(\Gamma_0, G)$.

(10.3) *The map h is a homeomorphism of $\mathfrak{R}_1(\Gamma_1', G_1) \times \mathfrak{R}_2$ onto $\mathfrak{R}_1(\Gamma_0, G)$.*

Proof. The continuity of h is obvious. To prove (10.3), it suffices to construct a map which is inverse to h. Suppose $r \in \mathfrak{R}_1(\Gamma_0, G)$. From (5.3) and (7.3), we know that $\pi_1 r(\Gamma_0)$ is a lattice in G_1 and that the kernel of $\pi_1 r(\Gamma_0)$ is $\Gamma_0 \cap G_2$. Therefore $\pi_1 r$ gives an isomorphism $p \colon \Gamma_1' \to G_1$, and moreover $p(\Gamma_1') = \pi_1 r(\Gamma_0)$ is a lattice in G_1. The map $r \to p$ is continuous. Now we define $u \colon \Gamma_0 \to G$ by $u(\gamma) = (\pi_1(\gamma), \pi_2 r(\gamma))$, $\gamma \in \Gamma_0$. It can be verified easily that u is an isomorphism and $u(\Gamma_0)$ is a lattice in G with $\gamma^{-1}u(\gamma) \in G_2$, $\gamma \in \Gamma_0$. This u depends continuously on r, and so u describes a connected set as r runs through $\mathfrak{R}_1(\Gamma_0, G)$. Hence $u \in \mathfrak{R}_1(\Gamma_0, G)$ and then $u \in \mathfrak{R}_2$. We have thus a continuous map $j \colon \mathfrak{R}_1(\Gamma_0, G) \to \mathfrak{R}_1(\Gamma_1', G_1) \times \mathfrak{R}_2$ given by $r \to (p, u)$. It is a trivial matter to see that both hj and jh are identity maps. (10.3) is thus proved.

Since G_1 is semi-simple and has no compact factor, the space $\mathfrak{R}_1(\Gamma_1, G)$ is known [15]. Therefore the study of $\mathfrak{R}_1(\Gamma_0, G)$ is reduced to the study of $\mathfrak{R}_2$. For this purpose, let $\mathfrak{R}^* = \{r \in \mathfrak{R}_1(\Gamma_0, G) : \gamma^{-1}r(\gamma) \in N, \gamma \in \Gamma_0\}$, and let $\mathrm{Ad}_G S$ denote the group of inner automorphisms of G induced by elements of S. This group $\mathrm{Ad}_G S$ is locally isomorphic with S and has a finite center.

Suppose $(\tau, r^*) \in \mathrm{Ad}_G S \times \mathfrak{R}^*$. We define $r \colon \Gamma_0 \to G$ by putting $r(\gamma) = \tau(r^*(\gamma))$, $\gamma \in \Gamma_0$. This r is evidently an element of $\mathfrak{R}_1(\Gamma_0, G)$. Let us see that r belongs to $\mathfrak{R}_2$. Taking account of the fact that

$$\pi_1(g) = \pi_1(\tau(g)), g \in G,$$

we have

$$\pi_1(\gamma^{-1}r(\gamma)) = \pi_1(\gamma^{-1}\tau r^*(\gamma)) = \pi_1(\gamma^{-1})\pi_1(r^*(\gamma))$$
$$= \pi_1(\gamma^{-1}r^*(\gamma)) \in \pi_1(N) = e.$$

This tells us that $\gamma^{-1}r(\gamma) \in G_2$, $\gamma \in \Gamma_0$, or what is the same, $r \in \mathcal{R}_2$. Thus we get a map $k\colon \mathrm{Ad}_G S \times \mathcal{R}^* \to \mathcal{R}_2$ defined by $(\tau, r^*) \to r$.

(10.4) *The map $k\colon \mathrm{Ad}_G S \times \mathcal{R}^* \to \mathcal{R}_2$ is a surjective local homeomorphism.*

Proof. The continuity of k is obvious. To see that it is surjective, let r be an arbitrary element of $\mathcal{R}_2$. From (9.6), there exists an element s of S such that

$$\lambda r(\gamma) = s\lambda(\gamma)s^{-1}, \quad \gamma \in \Gamma_0. \tag{10.5}$$

Define $q^*\colon \Gamma_0 \to G$ by $q^*(\gamma) = s^{-1}r(\gamma)s$, $\gamma \in \Gamma_0$. By (10.3) $\mathcal{R}_2$ is connected whence we see immediately that $q^* \in \mathcal{R}_1(\Gamma_0, G)$. Moreover, from (8.3), (10.5), and the definition of $\mathcal{R}_2$, we have

$$\pi(\gamma^{-1}q^*(\gamma)) = e, \ \lambda(\gamma^{-1}q^*(\gamma)) = \lambda(\gamma^{-1})\lambda(s^{-1}r(\gamma)s) = e,$$
$$\pi_1(\gamma^{-1}q^*(\gamma)) = \pi_1(\gamma^{-1}s^{-1}r(\gamma)s) = \pi_1(\gamma^{-1}s^{-1}\gamma)\pi_1(\gamma^{-1}r(\gamma))\pi_1(s) = e.$$

These three equalities tell us that $\gamma^{-1}q^*(\gamma) \in N$, or what is the same, $q^* \in \mathcal{R}^*$. Let σ be the inner automorphism of G induced by s. It is evident that $k(\sigma, q^*) = r$, and hence k is surjective.

Now let us show that k is locally one-to-one. Suppose that two elements (σ, r^*), (τ, q^*) of $\mathrm{Ad}_G S \times \mathcal{R}^*$ have the same image under k. Then $\sigma(r^*(\gamma)) = \tau(q^*(\gamma))$, or what is the same, $\tau^{-1}\sigma(r^*(\gamma)) = q^*(\gamma)$, $\gamma \in \Gamma_0$. Since both $\gamma^{-1}r^*(\gamma)$ and $\gamma^{-1}q^*(\gamma)$ belong to N, we have $\lambda r^*(\gamma) = \lambda(\gamma)$, $\lambda q^*(\gamma) = \lambda(\gamma)$. It follows then $\tau^{-1}\sigma(\lambda(\gamma)) = \lambda(\gamma)$, $\gamma \in \Gamma_0$. But $\lambda(\Gamma_0)$ is a lattice in S, and so the automorphism $\tau^{-1}\sigma$ must be the identity on S. Therefore $\tau^{-1}\sigma$ belongs to the center of $\mathrm{Ad}_G S$. From the semi-simplicity of $\mathrm{Ad}_G S$, we know then that k is locally one-to-one.

It remains for us to show that k is an open map. For this purpose, we fix a point (τ_1, r_1^*) in $\mathrm{Ad}_G S \times R^*$, and let $k(\tau, r_1^*) = r_1$. Choose a neighborhood U of τ_1 in $\mathrm{Ad}_G S$ such that, restricted to U, the covering map $f\colon \mathrm{Ad}_G S \to \mathrm{Ad}\, S$ is a homeomorphism. From (9.6), there exists a continuous map $b\colon \mathcal{R}_1(\Gamma_0, G) \to \mathrm{Ad}\, S$ with the property that

$$\lambda r(\gamma) = b(r)(\lambda(\gamma)), \gamma \in \Gamma_0, r \in \mathcal{R}_1(\Gamma_0, G).$$

Since $f(\tau_1) = b(r_1)$, we can find a neighborhood W of r_1 in $\mathfrak{R}_2$ such that $b(W) \subset f(U)$. For each $r \in W$, put $\tau = (f|_U)^{-1}b(r)$, $r^*(\gamma) = \tau^{-1}(r(\gamma))$, $\gamma \in \Gamma_0$. Then $r^* \in \mathfrak{R}^*$, and the map $r \to (\tau, r^*)$ is a well-defined map of W into $\mathrm{Ad}_G S \times \mathfrak{R}^*$. This map is a local inverse of k, and hence k is a local homeomorphism. This completes the proof.

From the above proof of (10.4), we know that if two elements (σ, r^*), (τ, q^*) of $\mathrm{Ad}_G S \times \mathfrak{R}^*$ have the same image under k, then $\tau^{-1}\sigma$ is an element, say ω, of the center F of $\mathrm{Ad}_G S$. It follows then that $q^*(\gamma) = \omega(r^*(\gamma))$, $\gamma \in \Gamma_0$. This suggests to us that k should be a regular covering map with F as the covering group. Now let us verify this in detail. Since $\mathrm{Ad}_G S$ is semi-simple and has a faithful linear representation, F is a finite group. To each $\omega \in F$, we define a transformation $T(\omega)$ of $\mathrm{ad}_G S \times \mathfrak{R}^*$ by putting $T(\omega)(\sigma, r^*) = (\sigma\omega^{-1}, \omega \circ r^*)$ where $\omega \circ r^*$ denotes the iteration of the two maps $r^*: \Gamma_0 \to G$, $\omega: G \to G$. It can be easily seen that

$$\omega \circ r^* \in \mathfrak{R}^*, T(\omega_1\omega_2) = T(\omega_1)T(\omega_2), kT(\omega) = \mathrm{k}, \ \omega, \omega_1, \omega_2 \in F$$

and thus $T(F)$ forms a finite transformation group of $\mathrm{Ad}_G S \times \mathfrak{R}^*$. If ω is not the identity, then $T(\omega)$ has no fixed point. Moreover, from the remarks at the beginning of this paragraph, two elements in $\mathrm{Ad}_G S \times \mathfrak{R}^*$ have the same image under k if and only if they belong to the same orbit under $T(F)$. Therefore

(10.5) *The map k: $\mathrm{Ad}_G S \times \mathfrak{R}^* \to \mathfrak{R}_2$ is a regular covering map with $T(F)$ as the covering group.*

Let $\mathfrak{R}_0^*$ be the connected component of $\mathfrak{R}^*$ containing the identity map r_0, and F_0 the subgroup of F consisting of all the elements ω such that $T(\omega)(\mathrm{Ad}_G S \times \mathfrak{R}_0^*) = \mathrm{Ad}_G S \times \mathfrak{R}_0^*$. Since $\mathfrak{R}^*$ is a closed subset of $\mathfrak{R}_1(\Gamma_0, G)$ defined by real analytic equations, $\mathfrak{R}^*$ is locally connected. From the connectedness of $\mathfrak{R}_2^*$, it follows then that the covering map k must carry $\mathrm{Ad}_G S \times \mathfrak{R}_0^*$ onto the entire $\mathfrak{R}_2$. We have therefore

(10.6) *$\mathrm{Ad}_G S \times \mathfrak{R}_0^*$ is a regular covering space of $\mathfrak{R}_2$ with k as the projection and the finite abelian group $T(F_0)$ as the covering group.*

11. **Main results.** From (10.3) and (10.6), the problem of deformations of Γ_0 in G is reduced to the study of the subspace $\mathfrak{R}_0^*$ of $\mathfrak{R}_1(\Gamma_0, G)$. To describe $\mathfrak{R}_0^*$, let us consider the group $\Gamma_0 N$. On account of (5.3), $\Gamma_0 N$ is closed in G and hence it is a Lie group (in general, not connected) with N as its identity component. Let A be the group of all automorphisms of $\Gamma_0 N$ leaving invariant each connected component of $\Gamma_0 N$. With respect to the

compact-open topology, A forms a topological group. This group is closely related with the space Q, with the compact-open topology, of all isomorphisms $r\colon \Gamma_0 \to G$ such that $r(\Gamma_0)$ are lattices in G and $\gamma^{-1}r(\gamma) \in N$ for all $\gamma \in \Gamma_0$. We note that $\mathcal{R}^* = Q \cap \mathcal{R}_1(\Gamma_0, G)$, and that $\mathcal{R}_0{}^*$ is the connected component of Q which contains the identity map $r_0\colon \Gamma_0 \to G$.

Let $a \in A$. Then $a(\Gamma_0)$ is obviously a lattice in G. Denoting, by $j(a)$, the restriction of a on Γ_0, we have therefore a map $j\colon A \to Q$.

(11.1) *The map j is a homeomorphism of A onto Q.*

Proof. To see that j is one-to-one, let $a, a' \in A$ and $j(a) = j(a')$. Then $a(\gamma) = a'(\gamma)$, $\gamma \in \Gamma_0$. Since Γ_0 is a lattice in G, $\Gamma_0 \cap N$ is a lattice in N. Now the two automorphisms a, a' being identical on the lattices $\Gamma_0 \cap N$ in N, they must be identical on the entire N. It follows that $a = a'$, and hence j is one-to-one.

To see that j is surjective, let us take $r \in Q$. Since $\gamma^{-1}r(\gamma) \in N$, $\gamma \in \Gamma_0$, we know that $r(\Gamma_0 \cap N) = r(\Gamma_0) \cap N$. The two lattices $\Gamma_0 \cap N$, $r(\Gamma_0) \cap N$ in N are therefore isomorphic. It follows then from Malcev's results [9] that there exists an automorphism σ of N such that

$$r(\beta) = \sigma(\beta), \quad \beta \in \Gamma_0 \cap N. \tag{11.2}$$

Hence

$$\sigma(\gamma\beta\gamma^{-1}) = r(\gamma\beta\gamma^{-1}) = r(\gamma)r(\beta)r(\gamma)^{-1} = r(\gamma)\sigma(\beta)r(\gamma)^{-1}$$
$$\beta \in \Gamma_0 \cap N, \quad \gamma \in \Gamma_0.$$

This means that the two automorphisms $\mathrm{Ad}_N r(\gamma) \circ \sigma$ and $\sigma \cdot \mathrm{Ad}_N \gamma$ of N coincide on the lattice $\Gamma_0 \cap N$ in N, and so they must coincide on the entire N. In other words,

$$\sigma(\gamma n \gamma^{-1}) = r(\gamma)\sigma(n)r(\gamma)^{-1},\ \gamma \in \Gamma_0,\ n \in N. \tag{11.3}$$

Now let us extend σ to an automorphism of $\Gamma_0 N$. Each element x of $\Gamma_0 N$ takes the form $x = \gamma n$, $\gamma \in \Gamma_0$, $n \in N$. On account of (11.2), if $\gamma_1 n_1 = \gamma_2 n_2$ with $\gamma_1, \gamma_2 \in \Gamma_0$, $n_1, n_2 \in N$, then $r(\gamma_1)\sigma(n_1) = r(\gamma_2)\sigma(n_2)$. Hence the map $a\colon \Gamma_0 N \to \Gamma_0 N$ defined by $a(x) = r(\gamma)\sigma(n)$ is a well-defined continuous map. Moreover, we can verify directly by using (11.3) that a is an automorphism of $\Gamma_0 N$. Since $\gamma^{-1}r(\gamma) \in N$, $\gamma \in \Gamma_0$, this automorphism a leaves invariant every connected component of $\Gamma_0 N$, and therefore $a \in A$. It is evident that $j(a) = r$. We have thus proved that j is surjective.

Now let us show that j is a homeomorphism. Suppose that $\{a_j\}$ be a sequence in A. We note that the convergence of $\{a_j\}$ in A is the uniform convergence on compact subsets of $\Gamma_0 N$ whereas the convergence of $\{j(a_j)\}$

is the uniform convergence on compact subsets of Γ. Therefore j is continuous. To see that j is open, let us assume that $\lim j(a_i) = j(a)$, where a is an element of A. Since N is a connected and simply-connected nilpotent Lie group, we can identify it with its Lie algebra $\hat{N}$ by means of the exponential map, and thus N has the structure of a real linear space. Each automorphism of N is then a linear transformation. According to Malcev [9], there exist elements $\beta_1, \beta_2, \cdots, \beta_m$ of the lattice $\Gamma_0 \cap N$ in N such that they form a base of the linear space N. Each automorphism of $\Gamma_0 N$ induces an automorphism of N, and therefore corresponds to a matrix with respect to the base $\beta_1, \beta_2, \cdots, \beta_m$ of N. Let M be the matrix corresponding to a, and M_i corresponding to a_i $(i = 1, 2, \cdots)$. Since $\lim j(a_i) = j(a)$, we have $\lim M_i = M$. This implies that, restricted to any compact subset of N, the sequence $\{a_i\}$ is uniformly convergent. Now let K be a compact subset of $\Gamma_0 N$. There exist elements $\gamma_1, \gamma_2, \cdots, \gamma_s$ of Γ_0 and compact subset K' of N such that

$$K \subset \gamma_1 K' \cup \gamma_2 K' \cup \cdots \cup \gamma_s K'.$$

Since $\{a_i\}$ is uniformly convergent on K' and also on the finite subset $\{\gamma_1, \gamma_2, \cdots, \gamma_s\}$ of Γ_0, it follows that $\{a_i\}$ is uniformly convergent on K. In other words, $\lim a_i = a$. Therefore j is a homeomorphism and (11.1) is proved.

Let A_0 be the e-component of the group of all automorphisms of $\Gamma_0 N$. Since N is the e-component of $\Gamma_0 N$, A_0 belongs to A and hence is the e-component of A. From (11.1), it follows directly

(11.4) *The map j carries A_0 homeomorphically onto $\mathfrak{R}_0{}^*$.*

Combining (10.3), (10.6) and (11.4), we have the following

MAIN THEOREM. *Suppose that G, Γ_0, G_1, G_2, S, N, π_1, π_2 have the same meaning as before. Let A_0 denote the e-component of the group of automorphisms of $\Gamma_0 N$ with the compact-open topology. For any $r_1 \in \mathfrak{R}_1(\pi_1(\Gamma_0), G_1)$, $\sigma \in \mathrm{Ad}_G S$ and $a \in A_0$, the map $r\colon \Gamma \to G$ defined by $r(\gamma) = (r_1\pi_1(\gamma), \pi_2\sigma a(\gamma))$, $\gamma \in \Gamma_0$ is an element of $\mathfrak{R}_1(\Gamma_0, G)$. The correspondence $(r_1, \sigma, a) \to r$ so obtained is a regular covering map of the product $\mathfrak{R}_1(\pi_1(\Gamma_0), G_1) \times \mathrm{Ad}_G S \times A_0$ onto $\mathfrak{R}_1(\Gamma_0, G)$ where the covering group is isomorphic with a subgroup of the center of $\mathrm{Ad}_G S$.*

This theorem reduces the problem of deformations of Γ_0 in G to the problem of deformations of lattices in G_1 which has been completely solved by A. Weil [15]. As an immediate consequence, we have

COROLLARY. *Under the same assumptions as above, the space* $\mathcal{R}_1(\Gamma_0, G)$ *is a manifold.*

Appendix.

Now we give a complete proof of the somewhat generalized version of the results of Auslander [2] and Zassenhaus [16] mentioned in Section 3. Since we do not assume the subgroup L to be discrete, the statement in Theorems A and B is more general, and hence some refinement of the proofs of these authors is required.

Let M be a subset of a topological group G. We shall use $\mathcal{L}_s(M)$ to denote the totality of commutatosr

$$[g_1[g_2[\cdot\cdot\cdot[g_{s-1}, g_s]\cdot\cdot\cdot]]], \quad g_i \in M$$

of length s. By $\lim_{s\to\infty} \mathcal{L}_s(M) = e$, we shall mean that given any neighborhood U of e, there exists an integer n such that $\mathcal{L}_s(M) \subset U$ for all $s \geqq n$.

LEMMA 1. *Let V be a normal subgroup of a connected Lie group G, and $\pi\colon G \to G/V$ the projection. If V is isomorphic with a real vector space, then there exists a neighborhood W of e in G/V such that, for any compact subset K of G with $\pi(K) \subset W$, $\lim_{s\to\infty} \mathcal{L}_s(K) = e$.*

This has been proved by L. Auslander in [1] and is also a direct consequence of formula (13) in [16] together with the Ado Theorem which asserts the existence of a local isomorphism ρ of G into $GL(R, n)$ such that $\rho(V)$ contains only unipotent matrices.

LEMMA 2. *Suppose that G, V, π, W have the same meaning as in Lemma 1. Let L be a closed subgroup of G with its e-component L_0 solvable. Then $L \cap \pi^{-1}(W)$ generates a solvable group.*

Proof. Let $K_1, K_2, \cdot\cdot\cdot$ be an increasing sequence of compact sets in G such that $\cup K_i = \pi^{-1}(W)$. Put $Q_i = K_i \cap L$, and *denote by P_i the subgroup generated by Q_i. From our choice of W,* $\lim_{s\to\infty} \mathcal{L}_s(Q_i) = e$ and hence there exists n such that $\mathcal{L}_n(Q_i) \subset L$. Let $\phi\colon L \to L/L_0$ be the projection. Then

$$\mathcal{L}_n(\phi(Q_i)) = \phi\mathcal{L}_n(Q_i) = e$$

and then $\phi(Q_i)$ generates a nilpotent group [16, p. 291]. Since L_0 is solvable and $\phi(P_i)$ is generated by $\phi(Q_i)$, P_i must be solvable. We have therefore an increasing sequence $P_1 \subset P_2 \subset P_3 \subset \cdot\cdot\cdot$ of solvable subgroups of G. Their union $P = \cup P_i$ is solvable [16] (Zassenhaus proved this when P_i are subgroups of $GL(R, m)$. But this can be trivially extended to the case that P_i

are subgroups of a connected Lie group by using the adjoint representation.) From the fact that $\cup K_i = \pi^{-1}(W)$, we see immediately that P is generated by $\pi^{-1}(W) \cap L$. The lemma is thus proved.

THEOREM A (Zassenhaus, Auslander). *Let G be a connected Lie group with its radical R simply-connected, $\pi\colon G \to G/R$ the projection, and L a closed subgroup of G. If the e-component L_0 of L is solvable, then the e-component of the closure of $\pi(L)$ is solvable.*

Proof. Let q be the length of the derived series of R. First we assume $q = 1$. By Lemma 2, there exists a neighborhood W of e in G/R such that $\pi^{-1}(W) \cap L$ generates a solvable group. Therefore the group H generated by $W \cap \pi(L)$ is solvable. Since W is open, $\pi(L) \cap W$ is dense in $\overline{\pi(L)} \cap W$, and hence $\overline{\pi(L)} \cap W \subset \bar{H}$. But the e-component $(\overline{\pi(L)})_0$ of $\overline{\pi(L)}$ is generated by any neighborhood of e in $\overline{\pi(L)}$. Therefore $(\overline{\pi(L)})_0$ is contained in $\bar{H}$ and is solvable. Theorem A is thus proved for $q = 1$.

Now assume Theorem A to be valid when the length of the derived series of R is less than q. Let $G^* = G/[R,R]$ and $\eta\colon G \to G^*$, $\xi\colon G^* \to G^*/\eta(R)$ $(= G/R)$ be the projections. By induction hypothesis, $L^* = \overline{\eta(L)}$ is a closed subgroup of G^* with its e-component $L_0{}^*$ solvable. Since $\eta(R)$ is isomorphic with a vector space, we know from the preceding paragraph that $\overline{\xi(L^*)}$ has a solvable e-component. Since $\overline{\pi(L)} = \overline{\xi(L^*)}$, it follows that $(\overline{\pi(L)})_0$ is solvable. This completes the induction and proves our theorem.

THEOREM B (L. Auslander). *Let G be a connected and simply-connected Lie group, R its radical, C the maximal connected normal compact subgroup of a semi-simple part of G, and $\theta\colon G \to G/CR$ the projection. Suppose that L is a closed subgroup of G with its e-component solvable. If $\theta(L)$ has the property (S) in G/CR, then $\theta(L)$ is discrete.*

Proof. Let $\pi\colon G \to G/R$, $\phi\colon G/R \to G/CR$ be the natural projections. From Theorem A, the e-component $(\overline{\pi(L)})_0$ of $\overline{\pi(L)}$ is solvable. The projection ϕ has a compact kernel, and so ϕ must be a closed map. Hence $\phi(\overline{\pi(L)})$ is closed and has $\phi((\overline{\pi(L)})_0)$ as its e-component. Since $\theta = \phi\pi$, we can see easily that $\phi((\overline{\pi(L)})_0)$ is invariant under $\mathrm{Ad}\,\theta(L)$. But G/CR is semi-simple and has ano compact factor, and $\theta(L)$ has the property (S) in G/CR. It follows [4] then that $\phi((\overline{\pi(L)})_0)$ is a normal subgroup of G/CR whence $\phi((\overline{\pi(L)})_0) = e$. This tells us that $\phi(\overline{\pi(L)})$ is discrete. The subgroup $\theta(L)$, being contained in $\phi(\overline{\pi(L)})$, is therefore discrete.

NORTHWESTERN UNIVERSITY AND
INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] L. Auslander, "Bieberbach's theorems on space groups and discrete uniform subgroups of Lie groups I," *Annals of Mathematics*, vol. 71 (1960), pp. 579-590.

[2] ———, *On radicals of discrete subgroups of Lie groups with applications to locally affine space.* Mimeographed notes.

[3] A. Borel, "Groupes linéaires algébriques," *Annals of Mathematics*, vol. 64 (1956), pp. 20-82.

[4] ———, "Density properties of certain subgroups of semi-simple groups without compact components," *Annals of Mathematics*, vol. 72 (1960), pp. 179-188.

[5] A. Borel and Harish-Chandra, "Arithmetic subgroups of algebraic groups," *Bulletin of the American Mathematical Society*, vol. 67 (1961), pp. 579-583.

[6] E. Calabi and E. Vesentini, "On compact, locally symmetric Kähler manifolds," *Annals of Mathematics*, vol. 71 (1960), pp. 472-507.

[7] C. Chevalley, *Théorie des groupes de Lie, III.* 1955.

[8] G. Frobenius, "Ueber den Jordanschen Satz," *Situngsberichte der Kgl. Preussischen Akademie der Wissenschaften zu Berlin*, 1911, pp. 241-248.

[9] A. I. Malcev, "On a class of homogeneous spaces," American Mathematical Society Translation No. 39 (1951).

[10] Y. Matsushima, "On the discrete subgroups and homogeneous spaces of nilpotent Lie groups," *Nagoya Mathematical Journal*, vol. 2 (1951), pp. 95-110.

[11] D. Montgomery and L. Zippin, "Theorem on Lie groups," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 448-452.

[12] G. D. Mostow, "Factor spaces of solvable groups," *Annals of Mathematics*, vol. 60 (1954).

[13] A. Selberg, "On discontinuous groups in higher-dimensional symmetric spaces," Contributions to Function theory, Tata Institute, 1960, pp. 147-164.

[14] A. Weil, "On discrete subgroups of Lie groups," *Annals of Mathematics*, vol. 72 (1960), pp. 369-384.

[15] ———, "On discrete subgroups of Lie groups, II," *Annals of Mathematics*, vol. 75 (1962), pp. 578-602.

[16] H. Zassenhaus, "Beweis eines Satzes ueber diskrete Gruppen," *Abh. Math. Sem. Hansischer Univ.*, vol. 19 (1938), pp. 289-312.

PARTITIONS INTO ODD SUMMANDS.*

By Peter Hagis, Jr.

1. **Introduction.** 1. In an earlier paper [1] the author has investigated the problem of determining the number of partitions of a positive number n into summands which are congruent modulo p to elements of a set a or their negatives, where p is an odd prime, and $a = \{a_1, a_2, \cdots, a_r\}$ with $1 \leqq a_i \leqq p/2$. A convergent series for this partition function was obtained by using the circle dissection method of Rademacher [6]. The necessary transformation equations as well as estimates of the magnitude of certain exponential sums were obtained by using the procedures of Lehner [5]. Asymptotic formulas were also derived. In the present paper we wish to extend our study so as to include the remaining prime, $p = 2$. That is, we wish to find a convergent series and asymptotic formulas for $q(n)$, the number of partitions of n into positive odd summands (or equivalently, into distinct positive summands). The methods employed are essentially the same as in [1] and free use will be made of the results obtained in the earlier paper whenever they are applicable.

The study of $q(n)$ is, of course, not new. In 1940, Hua [2], relying heavily on results from the theory of modular functions, was the first to represent $q(n)$ as a convergent infinite series. More recently, Iseki [3, 4] has obtained a convergent series representation, as well as asymptotic formulas, for a more general partition function which yields the number of partitions of an integer n into summands which are relatively prime to a given integer $M \geqq 2$. The case $M = 2$ (or $M = 4$) obviously gives $q(n)$. The circle dissection method is used by Iseki, but both his procedure for obtaining the transformation equations and his technique for reducing the exponential sums to Kloosterman sums differ from those of Lehner. Indeed, Iseki remarks that Lehner's method does not seem to work if $(M, 6) > 1$. For $p = 3$ it is shown in [1] that Lehner's procedure *can* be applied. In the present paper it is shown that Lehner's method, with only a slight modification, can be applied to the case $p = 2$.

* Received July 30, 1962; revised March 22, 1963.

2. The transformation equations. 2. The generating function of $q(n)$ is given by

$$F(x)=\prod_{m=0}^{\infty}(1-x^{2m+1})^{-1}=1+\sum_{n=1}^{\infty}q(n)x^{n}$$

which is convergent in the interior of the unit circle. Since it is necessary to determine the behavior of $F(x)$ in the neighborhood of rational points on a circle concentric to the unit circle and of radius less than one, we take $x=\exp\{2\pi ih/k-2\pi z/k\}$ where $\mathfrak{R}(z)>0$, $(h,k)=1$, $0\leqq h<k$.

With $r=1$, $a=\{1\}$, we have from Section II in [1]

$$G_a(x)=\sum_{n=1}^{\infty}\sum_{m=0}^{\infty}n^{-1}x^{(pm+1)n}+\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}n^{-1}x^{(pm-1)n}$$

satisfies the equation

$$G_a(x)=G_b(x')-2\pi i\{(Az-B/z)/12pki-\sigma_a(h,k)/2\}$$

if $p\mid k$, and

$$G_a(x)=J_a(x'')-2\pi i\{(Az-1/z)/12pki+\log(2\sin\pi\alpha/p)/2\pi i \\ -t_a(h,k)/2\}$$

if $p\nmid k$.

Here, $x'=\exp\{2\pi ih'/k-2\pi/kz\}$ and $x''=\{2\pi iH'/k-2\pi/Kz\}$ with h' defined by $hh'\equiv-1\pmod{k}$, H' defined by $phH'\equiv-1\pmod{k}$, and $K=pk$. $b=\{b_1\}$ where $b_1\equiv\pm h\pmod{p}$ whichever yields $0<b_1\leqq p/2$. $A=p^2-6p+6$, $B=p^2-6b_1p+6b_1^2$, and α is defined by $\alpha k\equiv1\pmod{p}$ with $0<\alpha<p$. Also,

$$\sigma_a(h,k)=\sum((\mu/k))((h\mu/k)),\mu=1,p+1,\cdots,k-p+1,p-1, \\ 2p-1,\cdots,k-1;$$

$$t_a(h,k)=\sum((\mu/K))((h\mu/k)),\mu=1,p+1,\cdots,K-p+1,p-1, \\ 2p-1,\cdots,K-1;$$

with $((x))=x-[x]-\tfrac{1}{2}+\tfrac{1}{2}\delta(x)$.

Finally,

$$J_a(x)=\sum_{m,n=1}^{\infty}m^{-1}\rho^{m}x^{mn}+\sum_{m,n=1}^{\infty}m^{-1}\bar{\rho}^{m}x^{mn}\quad\text{where}$$

$\rho=\exp(2\pi i\alpha/p)$, $\bar{\rho}=\exp(-2\pi i\alpha/p)$.

The proof of this result given in [1] for $p\geqq3$ also holds for $p=2$, in which case we have $b=\{1\}=a$, $A=B=-2$, $\alpha=1$. With $p=2$

$$\log F(x) = \tfrac{1}{2}\left(\sum_{n=1}^{\infty}\sum_{m=0}^{\infty} n^{-1}x^{(2m+1)n} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} n^{-1}x^{(2m-1)n}\right)$$
$$= \tfrac{1}{2}G_a(x).$$

By exponentiation we obtain

THEOREM 1. *$F(x)$ satisfies the transformation equation*

$$F(\exp\{2\pi ih/k - 2\pi z/k\}) = \omega(h,k)\exp\{\pi(z-1/z)/12k\} \cdot F(\exp\{2\pi ih'/k - 2\pi/kz\})$$

if $2 \mid k$, *and*

$$F(\exp\{2\pi ih/k - 2\pi z/k\}) = 2^{-\frac{1}{2}}\chi(h,k)\exp\{\pi(1/z+2z)/24k\} \cdot H(\exp\{2\pi iH'/k - \pi/kz\})$$

if $2 \nmid k$.

Here, $2hH' \equiv -1 \pmod{k}$;

(2.1) $$\omega(h,k) = \exp\{\pi i\sigma(h,k)\},\quad \sigma(h,k) = \sum((\mu/k))((h\mu/k)), \mu = 1,3,\cdots,k-1;$$

(2.2) $$\chi(h,k) = \exp\{\pi it(h,k)\}\quad t(h,k) = \sum((\mu/2k))((h\mu/k)), \mu = 1,3,\cdots,2k-1;$$

$$H(x) = \prod_{n=1}^{\infty}(1+x^n)^{-1} = F^{-1}(x).$$

This result is in agreement with Theorem 1 in [4] if we take $M = 2$ in the latter.

3. Estimates of certain exponential sums. 3. In the sequel it will be necessary to have some information concerning the magnitude of certain exponential sums involving $\omega(h,k)$ and $\chi(h,k)$. The trivial estimate $O(k)$ will not suffice so we must now undertake a study of these sums. Our procedure will be to reduce them to Kloosterman sums. The method used is essentially Lehner's. Many of the details will be omitted and the interested reader is referred to [1] and [5].

We must discuss three cases, $k \equiv 0 \pmod 4$, $k \equiv 2 \pmod 4$, and $k \equiv \pm 1 \pmod 4$. Considering first the case $4 \mid k$ we find (see (5.1) in [1] with $p = 4$, $r = 1$) that

$$24k\sigma(h,k) = 2h(2k^2 - 6k - 2) - 3k(k - 6 + 2c) - 48\sum_M \mu[h\mu/k]$$

where $M = \{\mu \mid \mu = 1, 5, \cdots, k-3\}$ and

$$c = \begin{cases} 1 \text{ if } h \equiv 1 \pmod 4 \\ 3 \text{ if } h \equiv 3 \pmod 4. \end{cases}$$

Thus, $24k\sigma(h,k)$ is always an integer. Furthermore,

(3.1) $$24k\sigma(h,k) \equiv 0 \pmod 3 \text{ if } 3 \nmid k.$$

We also have (see (5.5) in [1])

(3.2) $$24hk\sigma(h,k) = h^2(2k^2-6k-2)-(2+k^2)-6hk(c-2)-48kS$$

where $S = \frac{1}{2}\sum_M [h\mu/k]([h\mu/k]+1)$ is an integer.

Now let $fG = 48$ where f is the greatest divisor of 48 prime to k. If $3 \mid k$ then $f=1$, $G=48$, while if $3 \nmid k$ then $f=3$, $G=16$. If we now take h' so that $hh' \equiv -1 \pmod{Gk}$ and multiply (3.2) by h' we obtain

(3.3) $$24k\sigma(h,k) \equiv hu + h'v - 6k(c-2) \pmod{Gk}$$

where $u = 2k^2-6k-2$, $v = 2+k^2$.

Using (3.1) we see that

(3.4) $$24k\sigma(h,k) \equiv 0 \pmod f.$$

If we define the integer ϕ by the congruence $f\phi \equiv 1 \pmod{Gk}$ we have by (3.3) and (3.4)

$$24k\sigma(h,k) \equiv f\phi(uh+vh'-6k(c-2)) \pmod{Gkf}.$$

Therefore,

(3.5) $$\begin{aligned}\omega(h,k) &= \exp\{2\pi i(24k\sigma(h,k))/48k\} \\ &= \exp\{2\pi i(\phi(uh+vh')/Gk - 6\phi(c-2)/G)\}.\end{aligned}$$

4. We turn now to the case $k \equiv 2 \pmod 4$. Since Lehner's method must be modified slightly we shall be more explicit in our presentation here. We note first that from (2.1) and the properties of $((x))$ it is easily established that

$$\sigma(h,k) = \sum_P ((\mu/K))((h\mu/k))$$

where $K = 2k$ and $P = \{\mu \mid \mu = 1, 3, \cdots, K-1\}$. Thus,

(4.1) $$\sigma(h,k) = \sum_P (\mu/K - 1/2)((h\mu/k)) = \sum_P \mu K^{-1}((h\mu/k))$$

since $\sum_P ((h\mu/k)) = 0$. Writing $M = \{\mu \mid \mu = 1, 5, \cdots, K-3\}$ we have

$$\begin{aligned}\sigma(h,k) &= \sum_M \mu K^{-1}((h\mu/k)) + \sum_M (K-\mu)K^{-1}((h(K-\mu)/k)) \\ &= 2\sum_M \mu K^{-1}((h\mu/k))\end{aligned}$$

$$= \sum_M \mu k^{-1}(h\mu/k - [h\mu/k] - 1/2)$$

where we have used the fact that $\sum_M ((h\mu/k)) = \frac{1}{2}\sum_P ((h\mu/k)) = 0$. An easy calculation gives $hk^{-2}\sum_M \mu^2 = h(4k^2 - 6k - 1)/6k$ and $-\frac{1}{2}k^{-1}\sum_M \mu = -(k-1)/4$. Therefore,

$$12k\sigma(h,k) = 2h(4k^2 - 6k - 1) - 3k(k-1) - 12\sum_M \mu[h\mu/k]. \tag{4.2}$$

We conclude that $12k\sigma(h,k)$ is always an integer and that

$$12k\sigma(h,k) \equiv 0 \pmod 3 \text{ if } 3 \nmid k. \tag{4.3}$$

To obtain further information about $\sigma(h,k)$ we calculate $\sum_P ((h\mu/k))^2$ in two ways. In the first place

$$\begin{aligned} \sum_P ((h\mu/k))^2 &= \sum_P (h\mu/k - [h\mu/k] - 1/2)^2 \\ &= 2h\sum_P \mu k^{-1}(h\mu/k - [h\mu/k] - 1/2) - h^2k^{-2}\sum_P \mu^2 \\ &\quad + \sum_P [h\mu/k]([h\mu/k]+1) + \sum_P 1/4 \\ &= 4h\sigma(h,k) - 2h^2(K^2-1)/3K + 2S + k/4 \end{aligned} \tag{4.4}$$

where we have used (4.1) and $S = \frac{1}{2}\sum_P [h\mu/k]([h\mu/k]+1)$ is an integer.

Also, since $(h,k) = 1$ it follows that

$$\sum_P ((h\mu/k))^2 = \sum_P ((\mu/k))^2 = 2\sum_Q ((\mu/k))^2$$

where $Q = \{\mu \mid \mu = 1, 3, \cdots, k-1\}$.

Therefore,

$$\begin{aligned} \sum_P ((h\mu/k))^2 &= 2\sum_Q (\mu/k - 1/2)^2 \\ &= 2k^{-2}\sum_Q \mu^2 - 2k^{-1}\sum_Q \mu + \sum_Q 1/2 \\ &= (k^2-1)/3k - k/2 + k/4. \end{aligned} \tag{4.5}$$

Equating (4.4) and (4.5) we find that

$$48hk\sigma(h,k) = 4h^2(K^2-1) - (4+2k^2) - 24kS \tag{4.6}$$

Writing $k = 2k^*$ we now let $fG = 48$ where f is the greatest divisor of 48 prime to k^*. If $3 \mid k$ then $f = 16$ and $G = 3$, while if $3 \nmid k$ then $f = 48$ and $G = 1$. If we now define H^* by the congruence $4hH^* \equiv -1 \pmod{Gk^*}$, which is possible since $(4h, Gk^*) = 1$, and multiply (4.6) by H^* we have

$$(4.7) \qquad 12k\sigma(h,k) \equiv -h + 4H^* \pmod{Gk^*}.$$

Taking h' so that $hh' \equiv -1 \pmod{Gk}$ from which it follows that $hh' \equiv -1 \pmod{Gk^*}$ we have $4hh' \equiv -4 \pmod{Gk^*}$. Multiplying by H^* yields $h' \equiv 4H^* \pmod{Gk^*}$. Since $2 \nmid Gk^*$ there is an integer r such that $4r \equiv 1 \pmod{Gk^*}$. Therefore, $H^* \equiv rh' \pmod{Gk^*}$. From these remarks and (4.7) we have

$$(4.8) \qquad 12k\sigma(h,k) \equiv hu + h'v \pmod{Gk^*}$$

where $u = -1$, $v = 4r$.

A discussion of congruence modulo 16 is now necessary. Since $\mu \equiv 1 \pmod 4$ we have $12 \sum_M \mu[h\mu/k] \equiv 12 \sum_M [h\mu/k] \pmod{16}$. Now,

$$\begin{aligned} 12 \sum_M [h\mu/k] &= -12 \sum_M ((h\mu/k)) + 12hk^{-1} \sum_M \mu - 6 \sum_M 1 \\ &= 6h(k-1) - 3k \\ &= 12hk^* - 6h - 3k. \end{aligned}$$

Therefore, from (4.2) it follows that

$$12k\sigma(h,k) \equiv -12h(k+k^*) + 4h + 3(2k - k^2) \pmod{16}.$$

Since $h = 4n + d$ where $d = -3$ if $h \equiv 1 \pmod 4$ and $d = 3$ if $h \equiv 3 \pmod 4$ we have finally

$$(4.9) \qquad 12k\sigma(h,k) \equiv -12d(k+k^*) + 4d + 3(2k-k^2) \pmod{16}.$$

From (4.3) and (4.9) it follows that

$$(4.10) \qquad 12k\sigma(h,k) \equiv -12d(k+k^*) + 4d + 3(2k-k^2) \pmod{f}.$$

Defining ϕ and Γ by the congruences $f\phi \equiv 1 \pmod{Gk^*}$, $Gk^*\Gamma \equiv 1 \pmod f$ we have by (4.8) and (4.10),

$$12k\sigma(h,k) \equiv f\phi(uh + vh') + Gk^*\Gamma(-12d(k+k^*) + 4d + 3(2k-k^2)) \pmod{fGk^*}.$$

Therefore,

$$(4.11) \qquad \begin{aligned} \omega(h,k) &= \exp\{2\pi i(12k\sigma(h,k))/24k\} \\ &= \exp\{2\pi i(\phi(uh+vh')/Gk^* \\ &\qquad + \Gamma(-12d(k+k^*) + 4d + 3(2k-k^2))/f)\}. \end{aligned}$$

5. We now consider the case $2 \nmid k$. From (2.2) and the properties of $((x))$ it is easily established that $t(h,k) = \sum_P ((\mu/K))((h\mu/k))$ where $K = 4k$ and $P = \{\mu \mid \mu = 1, 3, 5, \cdots, K-1\}$. It then follows (see (6.1) in [1] with $p = 4$, $r = 1$) that

$$(5.1)\quad 24kt(h,k)=2h(2K^2-6K-2)-3k(K-2-2\alpha)-12\sum_M \mu[h\mu/k]$$

where $M=\{\mu \mid \mu=1,5,\cdots,K-3\}$ and α is defined by $k\alpha\equiv 1 \pmod 4$, $0<\alpha<4$. Thus, $24kt(h,k)$ is always an integer, and

$$(5.2)\qquad 24kt(h,k)\equiv 0 \pmod 3 \text{ if } 3\nmid k.$$

We also have (see (6.5) in [1])

$$(5.3)\qquad 24hkt(h,k)=h^2(2K^2-6K-2)+(1-k^2)+6hk\alpha-12kT$$

where $T=\frac{1}{2}\sum_M [h\mu/k]([h\mu/k]+1)$ is an integer.

Now let $Fg=48$ where F is the greatest divisor of 48 prime to k. If $3\mid k$ then $F=16$, $g=3$, while if $3\nmid k$ then $F=48$, $g=1$. If we take h' so that $hh'\equiv -1 \pmod{gk}$ and multiply (5.3) by h' we obtain

$$(5.4)\qquad 24kt(h,k)\equiv hu+h'v \pmod{gk}$$

where $u=-2$, $v=-1$.

A discussion of congruence modulo 16 is now necessary. Since $\mu\equiv 1 \pmod 4$ we have $12\sum_M \mu[h\mu/k]\equiv 12\sum_M [h\mu/k] \pmod{16}$. Also, $\sum_M ((h\mu/k))=0$ since the μ in M constitute a complete residue system modulo k. Therefore,

$$\begin{aligned}12\sum_M [h\mu/k] &= -12\sum_M ((h\mu/k))+12hk^{-1}\sum_M \mu-6\sum_M 1+6\sum_M \delta(h\mu/k)\\ &=24hk-12h-6k+6.\end{aligned}$$

From (5.1) it follows that

$$(5.5)\qquad 24kt(h,k)\equiv 12k(1-k)+6(k\alpha-1) \pmod{16}.$$

From (5.2) and (5.5) it follows that

$$(5.6)\qquad 24kt(h,k)\equiv 12k(1-k)+6(k\alpha-1) \pmod F.$$

Defining Φ and γ by the congruences $F\Phi\equiv 1 \pmod{gk}$, $gk\gamma\equiv 1 \pmod F$ we have, by (5.4) and (5.6),

$$24kt(h,k)\equiv F\Phi(uh+vh')+gk\gamma(12k(1-k)+6(k\alpha-1)) \pmod{Fgk}.$$

Therefore,

$$(5.7)\qquad \begin{aligned}\chi(h,k) &= \exp\{2\pi i(24kt(h,k))/48k\}\\ &=\exp\{2\pi i(\Phi(uh+vh')/gk\\ &\qquad +\gamma(12k(1-k)+6(k\alpha-1))/f)\}.\end{aligned}$$

6. THEOREM 2. *The sum*

$$A(n,\nu;k;d;\sigma_1,\sigma_2)=\sum_{h \bmod k}{}' \omega(h,k)\exp\{-2\pi i(hn-h'\nu)/k\}$$

where $h\equiv d \pmod 4$, $2\nmid d$; $\sigma_1 \leqq h' < \sigma_2 \pmod k$, $0\leqq \sigma_1<\sigma_2\leqq k$, $2\mid k$, *is subject to the estimate* $O(n^{1/3}k^{2/3+\epsilon})$ *uniformly in* $\nu, d, \sigma_1, \sigma_2$.

THEOREM 3. *The sum*

$$B(n,\nu;k;\sigma_1,\sigma_2)=\sum_{h \bmod k}{}' \chi(h,k)\exp\{-2\pi i(hn-H'\nu)/k\}$$

where $\sigma_1 \leqq h' < \sigma_2 \pmod k$, $0\leqq \sigma_1<\sigma_2\leqq k$, $2\nmid k$ *is subject to the estimate* $O(n^{1/3}k^{2/3+\epsilon})$ *uniformly in* ν, σ_1, σ_2.

Here σ_1, σ_2 are integers, $hh'\equiv -1 \pmod k$, $2hH'\equiv -1 \pmod k$ and $\sum'$ indicates that h runs over integers prime to the modulus of the sum. Using (3.5), (4.11) and (5.7) the proofs of these theorems are identical with those of Theorems 2 and 3 in [1] and are therefore omitted.

4. A convergent series and asymptotic formulas for $q(n)$. 7. By Cauchy's integral formula we have

$$q(n)=\frac{1}{2\pi i}\int_C x^{-n-1}F(x)\,dx=\sum_{h,k}{}'\frac{1}{2\pi i}\int_{\xi_{h,k}} x^{-n-1}F(x)\,dx$$

where $0\leqq h<k\leqq N$ and $\xi_{h,k}$ are the Farey arcs of order N of C, the circle $|x|=\exp\{-2\pi N^{-2}\}$. If on the arc $\xi_{h,k}$ we introduce the variable ϕ by the equation $x=\exp\{-2\pi N^{-2}+2\pi ih/k+2\pi i\phi\}$ and write $w=N^{-2}-i\phi$, $z=wk$ we obtain

$$q(n)=\sum_{h,k}{}'\exp\{-2\pi inh/k\}\int_{-\theta'}^{\theta''}F(\exp\{2\pi ih/k-2\pi z/k\})\exp\{2\pi nw\}\,d\phi.$$

Here $\theta'=1/k(k+k_1)$ and $\theta''=1/k(k+k_2)$ where $h_1/k_1<h/k<h_2/k_2$ are consecutive terms in the Farey series of order N.

Splitting the sum over k into 2 parts $q(n,0)$ and $q(n,1)$ according to whether $2\mid k$ or $2\nmid k$ respectively we have $q(n)=q(n,0)+q(n,1)$. By Theorem 1,

$$q(n,0)=\sum_{h,k}{}'\omega(h,k)\exp\{-2\pi inh/k\}\int_{-\theta'}^{\theta''}\sum_{\nu=0}^{\infty}q(\nu)\exp\{2\pi ih'\nu/k\}$$
$$\cdot\exp\{-(\pi/k^2w)(2\nu+1/12)+\pi w(2n+1/12)\}\,d\phi$$

where $2\mid k$.

If we split $q(n,0)$ into 2 parts according as $h\equiv 1 \pmod 4$ or $h\equiv 3$

(mod 4) then, since $2\nu + 1/12 > 0$, we have using Rademacher's argument [6] and Theorem 2

$$q(n,0) = 0(n^{1/8}N^{-1/8+\epsilon}\exp\{2\pi nN^{-2}\}).$$

Turning to $q(n,1)$ we have by Theorem 1

$$q(n,1) = 2^{-1/2}\sum_{h,k}{}'\chi(h,k)\exp\{-2\pi inh/k\}\int_{-\theta'}\sum_{\nu=0}^{\infty}c(\nu)\exp\{2\pi iH'\nu/k\}$$
$$\cdot\exp\{-(\pi/2k^2w)(2\nu-1/12)+\pi w(2n+1/12)\}d\phi$$

where $2\nmid k$ and where we have written $H(x)=\sum_{\nu=0}^{\infty}c(\nu)x^{\nu}$. Splitting the sum over ν into 2 parts according as $\nu=0$ or $\nu>0$ we have

$$q(n,1) = Q(n) + R(n).$$

Since $2\nu - 1/12 > 0$ if $\nu > 0$ we have, again using Rademacher's technique, and employing Theorem 3,

$$R(n) = 0(n^{1/8}N^{-1/8+\epsilon}\exp\{2\pi nN^{-2}\}).$$

For $Q(n)$ (see (8.8), (8.9), (8.10) in [1] with $p=4$, $r=1$) we have

$$Q(n) = \pi\sum_{k=1}^{N}B(k,n)L(k,n) + 0(n^{1/8}N^{-1/8+\epsilon}\exp\{2\pi nN^{-2}\})$$

where

(7.1) $$B(k,n) = \sum_{h \bmod k}{}'\chi(h,k)\exp\{-2\pi inh/k\},$$

and

(7.2) $$L(k,n) = k^{-1}(24n+1)^{-1/2}I_1\{\pi(48n+2)^{1/2}/12k\}.$$

Here $I_1(z)$ is the Bessel function of first order and $2\nmid k$.

Since $q(n) = q(n,0) + Q(n) + R(n)$ we have, letting $N\to\infty$,

THEOREM 4. *The number of partitions, $q(n)$, of an integer n into positive odd summands (or into distinct positive summands) is given by the convergent series*

(7.3) $$q(n) = \pi\sum_{k=1}^{\infty}B(k,n)L(k,n)$$

where $2\nmid k$ and $B(k,n)$, $L(k,n)$ are given by (7.1) and (7.2) respectively. This result is easily seen to agree with those of Hua [2] and Iseki [4].

Asymptotic formulas for $q(n)$ can be obtained by using the procedures detailed in [1]. Indeed, since (7.3) is obtainable formally from (8.11) in

[1] by setting $r=1$, $p=4$ in the latter, we have by Corollaries 7.1 and 7.2 in [1].

THEOREM 5. *As* $n\to\infty$

$$q(n)=\pi(24n+1)^{-1/2}I_1\{\pi(48n+2)^{1/2}/12\}(1+0(\exp\{-cn^{1/2}\}))$$

where $c>0$.

COROLLARY. *As* $n\to\infty$

$$q(n)=18^{1/4}(24n+1)^{-3/4}\exp\{\pi(48n+2)^{1/2}/12\}(1+0(n^{-1/2})).$$

These results are in agreement with those of Iseki [4].

TEMPLE UNIVERSITY, PHILADELPHIA, PENNSYLVANIA.

REFERENCES.

[1] P. Hagis, Jr., "A problem on partition with a prime modulus $p\geqq 3$," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 30-62.

[2] L. K. Hua, "On the number of partitions of a number into unequal parts," *Transactions of the American Mathematical Society*, vol. 51 (1942), pp. 194-201.

[3] S. Iseki, "A partition function with some congruence condition," *American Journal of Mathematics*, vol. 81 (1959), pp. 939-961.

[4] ———, "Partitions in certain arithmetic progressions," *American Journal of Mathematics*, vol. 83 (1961), pp. 243-264.

[5] J, Lehner, "A partition function connected with the modulus five," *Duke Mathematical Journal*, vol. 8 (1941), pp. 631-655.

[6] H. Rademacher, "The Fourier coefficients of the modular invariant $J(\tau)$, *American Journal of Mathematics*, vol. 60 (1938), pp. 501-512.

EXISTENCE OF UNIVERSAL CONNECTIONS II.*

By M. S. NARASIMHAN and S. RAMANAN.

1. Introduction. In an earlier paper [3] we proved the existence of universal connections for connections in bundles with a compact Lie group as structure group. In this paper we extend this result to the case of an arbitrary connected Lie group (Theorem 1). The proof of this theorem does not depend on [3]. However, this result does not include Theorem 1 of [3] which is more precise in that it asserts that the canonical connections in the Stiefel bundles themselves are universal for connections in unitary or orthogonal bundles. The latter is useful in some applications.

Since any two connections on a principal bundle differ by a 1-form of the adjoint type, one can reduce the problem of finding a universal connection to one of finding a universal 1-form of the adjoint type (§ 3). Regarding the latter problem we prove the following more general result (Theorem 2): if ρ is a finite dimensional representation of a connected Lie group G and n and p are non-negative integers, then there exists a n-universal p-form of type ρ. (For the definition of forms of type ρ see § 2.) This problem is essentially one for compact Lie groups (§ 6) since the structure group of a G-bundle P can be reduced to a maximal compact subgroup K of G, and forms of type ρ on P are precisely forms on P obtained by extending forms of type $\rho|K$ (restriction of ρ to K) on the reduced bundle (§ 2). It should be remarked, however, that the existence of universal connections for a connected Lie group does not follow immediately from that for compact Lie groups, since not every connection on a G-bundle is the extension of a connection on a reduced K-bundle. (For instance, the holonomy group of a connection got by extension will have to be contained in K). In the case of connections, Theorem 2 seems to be necessary for passing from the compact to the general case. In our procedure, however, Theorem 2 implies at once Theorem 1 without passing through the compact case.

In the case when G is the orthogonal group $O(k)$ and ρ is the natural representation, an n-universal p-form is constructed explicitly (§ 4). If G is compact we may suppose that ρ is a representation of G by orthogonal matrices. This enables one to reduce the compact to the orthogonal case (§ 5).

* Received May 7, 1962.

For the notions relating to connections in principal bundles we refer to [1], [2], [5].

2. Preliminaries. In this section, we first fix our notation and terminology, and then give a canonical way of extending p-forms of a certain type to bundles obtained by extension of structure group.

By 'differentiable' we always mean 'indefinitely differentiable.' All manifolds, groups, bundles, maps and forms are assumed differentiable. We assume all our manifolds are paracompact. By a p-form on a manifold we mean a covariant tensor of degree p. If $f\colon M \to M'$ is a map and α a p-form on M', $f^*(\alpha)$ will denote the inverse image of α by f.

By a G-bundle we always mean a *principal* bundle with structure group G. If $f\colon H \to G$ is a homomorphism of groups, P_1 a H-bundle, and P_2 a G-bundle, a map $h\colon P_1 \to P_2$ will be called a f-morphism if $h(\xi s) = h(\xi)f(s)$ for every $\xi \in P_1$, $s \in H$. When $H = G$ and f is the identity map, a f-morphism will be called a G-morphism. If ρ is a representation of G in a finite dimensional vector space V, a p-form on a G-bundle with values in V is said to be of type ρ if i) it is equivariant for the action of G and ii) it annihilates any p-tuple of tangent vectors one of which is vertical [2]. If $h\colon P_1 \to P_2$ is a f-morphism and α a p-form of type ρ on P_2, then $h^*\alpha$ is clearly a p-form on P_1 of type $(\rho \circ f)$.

Let G_1, G_2 be two Lie groups and $f\colon G_1 \to G_2$ a homomorphism. Let T_f be the functor which associates to every G_1-bundle P the G_2-bundle $T_f(P)$ over the same base obtained by extension of the structure group by f. We recall that the total space of $T_f(P)$ is the orbit space of $P \times G_2$ under the action of G_1 given by $(\xi, g_2)g_1 = (\xi g_1, f(g_1)^{-1}g_2)$ for $\xi \in P$, $g_1 \in G_1$, $g_2 \in G_2$. The action of G_2 on $P \times G_2$ defined by $(\xi, g_2)g_2' = (\xi, g_2 g_2')$ for $\xi \in P$, $g_2, g_2' \in G_2$ commutes with the above action of G_1 and hence G_2 operates on $T_f(P)$ and makes of it a G_2-bundle. Moreover, if $\Phi\colon P \to P'$ is a G_1-morphism, the G_2-morphism $T_f(\Phi)\colon T_f(P) \to T_f(P')$ is induced by the map $(\xi, g_2) \to (\Phi(\xi), g_2)$ of $P \times G_2$ into $P' \times G_2$.

Let q be the projection $P \times G_2 \to T_f(P)$ and i_f the map $\xi \to q(\xi, e)$ of P into $T_f(P)$. We then have $i_f(\xi s) = i_f(\xi)f(s)$, $\xi \in P$, $s \in G_1$; i.e., i_f is a f-morphism.

Let ρ be a finite dimensional representation of G_2. If β is a p-form of type ρ on $T_f(P)$, $i_f^*\beta$ is a p-form on P of type $\rho \circ f$. Conversely, to every p-form α on P of type $\rho \circ f$ we can associate in a natural way a p-form $T_f(\alpha)$ of type ρ on $T_f(P)$ with $i_f^*T_f(\alpha) = \alpha$ in the following way. It is easy to check that the form α' on $P \times G_2$ defined by $\alpha'_{(\xi, g_2)} = \rho(g_2)^{-1}(p_1^*\alpha)_{(\xi, g_2)}$

(p_1 being the projection $P \times G_2 \to P$) is invariant under the action of G_1 and annihilates any p-tuple of tangent vectors one of which is vertical. Hence there exists a p-form $T_f(\alpha)$ of type ρ on $T_f(P)$ such that $q^*T_f(\alpha) = \alpha'$ and we have $i_f^*T_f(\alpha) = \alpha$. Moreover, if β is a p-form of type ρ on $T_f(P)$, we have $T_f(i_f^*\beta) = \beta$.

Finally, the correspondence $\alpha \to T_f(\alpha)$ is 'functorial' in the following sense: if $\Phi: P \to P'$ is a G_1-morphism and if α' is a p-form on P' of type $\rho \circ f$, then we have

$$T_f(\Phi^*\alpha') = (T_f\Phi)^*(T_f\alpha').$$

3. Statement of the theorems. Proof of Theorem 1.

THEOREM 1. *Let G be a connected Lie group and n a positive integer. Then there exist a principal G-bundle B and a connection form γ_0 on B such that any connection form on a principal G-bundle P with base of dimension $\leqq n$ is the inverse image of γ_0 by a G-morphism of P in B.*

We deduce Theorem 1 from the following theorem which seems to be of independent interest.

THEOREM 2. *Let G be a connected Lie group and ρ a finite dimensional representation of G. Let n and p be two non-negative integers. Then there exist a principal G-bundle E and a p-form α_0 of type ρ on E such that any p-form of type ρ on a principal G-bundle P with base of dimension $\leqq n$ is the inverse image of α_0 by a G-morphism of P into E. Moreover, the bundle E can be chosen to be classifying for dimension $\leqq n$.*

Remarks. 1. A p-form (resp. a connection) which possesses the property stated in Theorem 2 (resp. Th. 1) will be called n-universal.

2. Theorem 2 is also valid with "p-form" replaced by "exterior p-form." A universal exterior p-form is obtained by alternating a universal p-form.

Proof of Theorem 1. We now prove Theorem 1 assuming Theorem 2. Let F be a differentiable G-bundle which is n-universal, and γ_1 any connection on F. On the other hand, let E be a G-bundle and α_0 a 1-form on E of the adjoint type which is n-universal for such forms. Consider the G-bundle $B = F \times E$, the action of G on B being given by $(f, e)g = (fg, eg)$, $f \in F$, $e \in E$, $g \in G$. Let $q_1: B \to F$, $q_2: B \to E$ be the canonical projections, which are clearly G-morphisms. The differential form $\gamma_0 = q_1^*\gamma_1 + q_2^*\alpha_0$ is a connection form on B since $q_1^*\gamma_1$ is a connection form and $q_2^*\alpha_0$ is a 1-form of the adjoint type. We assert that γ_0 is n-universal for connections in G-bundles.

In fact let P be any G-bundle with base of dimension $\leqq n$ and γ any connection on P. Since F is a n-universal bundle, there exists a G-morphism $\Psi_1\colon P\to F$. Then $\gamma-\Psi_1^*(\gamma_1)$ is a 1-form of the adjoint type since γ and $\Psi_1^*(\gamma_1)$ are connection forms on P. Let $\Psi_2\colon P\to E$ be a G-morphism such that $\Psi_2^*(\alpha_0)=\gamma-\Psi_1^*(\gamma_1)$. Consider the G-morphism $\Phi\colon P\to B$ defined by $q_1\circ\Phi=\Psi_1$; $q_2\circ\Phi=\Psi_2$. Then we have

$$\begin{aligned}\Phi^*(\gamma_0)&=\Phi^*(q_1^*\gamma_1+q_2^*\alpha_0)\\&=(q_1\circ\Phi)^*\gamma_1+(q_2\circ\Phi)^*\alpha_0\\&=\Psi_1^*(\gamma_1)+\Psi_2^*(\alpha_0)\\&=\Psi_1^*(\gamma_1)+(\gamma-\Psi_1^*(\gamma_1))\\&=\gamma.\end{aligned}$$

Remark. In the above construction, it is clear that the bundle B is n-classifying if the G-bundles E and F are n-classifying, so that the maps induced on the bases by two G-morphisms $P\to B$ are homotopic. Thus the theorem of A. Weil on connections [1] is an immediate consequence of Theorem 1. However, Weil's theorem can be proved in a simpler way; for, all one requires for the proof is that any *two* given connections γ_1 and γ_2 on a bundle P can be obtained as the inverse images of the *same* connection γ_0 on a bundle B by morphisms whose projections on the base are homotopic. This problem is considerably simpler as can be seen by taking $B=P\times I$ and $(\gamma_0)_{(\xi,t)}=tp^*(\gamma_2)+(1-t)p^*\gamma_1$, $\xi\in P$, $t\in I$ where p is the projection $P\times I\to P$ (I is the open interval $[-2,2]$). The inclusions $P\to P\times I$ given by $\xi\to(\xi,0)$ and $\xi\to(\xi,1)$ induce γ_1, γ_2 respectively and their projections to the base are clearly homotopic, the homotopy being induced by the identity mapping of $P\times I$ into itself [3, § 6].

4. The orthogonal case. In this section, we prove Theorem 2 in the case where G is the real orthogonal group $O(k)$ and ρ is the natural representation in $\boldsymbol{R}^k$. We identify $O(k)$ with the group of (k,k) real matrices A such that $A'A=I_k$ (A' being the transpose of A) and $\boldsymbol{R}^k$ with $(k,1)$ real matrices. ρ then corresponds to left multiplication of $(k,1)$ matrices by (k,k) orthogonal matrices.

Let $W(N,k)$, $N\geqq k$, be the Stiefel bundle of (N,k)-real matrices A such that $A'A=I_k$ ([3, § 2]). $O(k)$ acts on $W(N,k)$ by multiplication on the right. If $A\in W(N,k)$ is of the form $\begin{bmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_N\end{bmatrix}$ where each A_i is a $(1,k)$

matrix, the function σ_i on $W(N,k)$ which assigns to each A the $(k,1)$ matrix A_i' is of type ρ. For

$$\sigma_i(As)=\sigma_i\begin{bmatrix}A_1s\\ \vdots\\ A_Ns\end{bmatrix}=(A_is)'=s'A_i'=s^{-1}\sigma_i(A)$$

for $A\in W(N,k)$, $s\in O(k)$.

We now construct a n-universal p-form of type ρ. For the rest of this section, N will denote the integer $(n+1)n^p+(n+k)$. Let $V_1,\cdots,V_{n+1}$ be $(n+1)$ copies of $\boldsymbol{R}^n$ and $V_0=\boldsymbol{R}$. Consider the $O(k)$-bundle

$$E=W(N,k)\times V_0\times V_1\times\cdots\times V_{n+1}$$

where the action of $O(k)$ on E is given by

$$(w,v_0,\cdots,v_{n+1})g=(wg,v_0,\cdots,v_{n+1}),\ w\in W(N,k),\ v_i\in V_i,\ g\in O(k).$$

Let π (resp. π_i) denote the projection of E onto $W(N,k)$ (resp. V_i). Let further $(x_i^1,\cdots,x_i^n)$ be the coordinate functions in V_i, $i>0$. For each multi-index $I=(i_1,\cdots,i_r,\cdots,i_p)$, $1\leqq i_r\leqq n$ and $1\leqq j\leqq n+1$, we shall denote by ω_j^I the p-form $\pi_j^*(dx_j^{i_1}\otimes\cdots\otimes dx_j^{i_p})$ on E. For convenience of notation, let us choose a bijection λ of the set of indices (I,j) with $I=(i_1,\cdots,i_p)$ $1\leqq i_r\leqq n$ and $1\leqq j\leqq n+1$, onto the set of integers $[1,(n+1)n^p]$. Obviously $\tau_j^I=\sigma_{\lambda(I,j)}\circ\pi$ is a function on E with values in $(k,1)$-matrices of type ρ. π_0 being a real-valued function on E, the form $(\pi_0\cdot\tau_j^I)\omega_j^I$ is a $(k,1)$-matrix valued form which is the product of the vector valued function $\pi_0\tau_j^I$ and the real-valued form ω_j^I. The form

$$\alpha_0=\sum_{I,j}(\pi_0\tau_j^I)\omega_j^I$$

is of type ρ. For, clearly α_0 annihilates any p-tuple of vectors one of which is vertical since each ω_j^I has this property. Moreover if $X_1,\cdots,X_p$ are vectors at $\xi\in P$ and $s\in O(k)$ we have

$$\begin{aligned}\alpha_0(X_1s,\cdots,X_ps)&=\sum\pi_0(\xi s)\tau_j^I(\xi s)\omega_j^I(X_1s,\cdots,X_ps)\\&=s^{-1}\{\pi_0(\xi)\sum\tau_j^I(\xi)\omega_j^I(X_1,\cdots,X_p)\}\\&=s^{-1}\alpha_0(X_1,\cdots,X_p)\end{aligned}$$

where the X_is are the vectors at ξs which are images of X_i by the differential of the map $\xi\to\xi s$ of P into itself.

We now proceed to prove that α_0 is n-universal for p-forms of type ρ.

Proof. Let P be a $O(k)$-bundle over a base M of dimension $\leqq n$ and

$q: P \to M$ be the projection. Let (U_i) be a covering of M by relatively compact open sets such that

i) each $\bar{U}_i$ is contained in a coordinate cell

ii) the U_i's can be divided into $(n+1)$ classes $\mathcal{B}_j$ in such a way that no two U_i's of the same class intersect [4, p. 61].

Let W_i be a shrinking of this covering, i.e., an open covering W_i such that $\bar{W}_i \subset U_i$. Let $D_j\{j=1,\cdots,(n+1)\}$ be the union of the open sets $q^{-1}(W_i)$ for those i's for which U_i belongs to $\mathcal{B}_j$. Let ζ_j be a partition of unity with respect to this covering, consisting of non-negative differentiable functions ζ_j invariant under the action of G with support of $\zeta_j \subset D_j$ and $\sum \zeta_j = 1$.

Let α be a p-form of type ρ on P. It is clear that there exist functions $(f_j^1,\cdots,f_j^n)$ on M whose restrictions to each W_i, for those i's for which U_i belongs to $\mathcal{B}_j$, form a coordinate system on W_i. Since α is of type ρ, α can be expressed in D_j in the form $\sum \alpha_j{}^I q^*(df_j{}^I)$, where $\alpha_j{}^I$ are functions of type ρ on D_j and $df_j{}^I = df_j{}^{i_1} \otimes \cdots \otimes df_j{}^{i_p}$. Then it is easy to see that $\alpha = \sum \beta_j{}^I q^*(df_j{}^I)$, where $\beta_j{}^I = \zeta_j \alpha_j{}^I$ are now differentiable functions on P of type ρ. Let h be a strictly positive invariant differentiable function on P such that $h(\xi)^2 > 2\,\|\,\beta_j{}^I(\xi)\beta_j{}^I(\xi)'\,\|$ for every $\xi \in P$, where $\|\ \|$ denotes the norm as a linear operator. (The existence of such an h follows for instance from the fact that $\|\,\beta_j{}^I(\xi)\beta_j{}^I(\xi)'\,\|$ is an invariant function on P). We have

$$\alpha = \sum h\eta_j{}^I q^*(df_j{}^I)$$

where $\eta_j{}^I = \dfrac{1}{h}\beta_j{}^I$. Obviously

$$\|\,\eta_j{}^I(\xi)\eta_j{}^I(\xi)'\,\| = \Big\|\,\frac{1}{h(\xi)^2}\sum \beta_j{}^I(\xi)\beta_j{}^I(\xi)'\,\Big\| \leqq \tfrac{1}{2}.$$

Therefore $R(\xi) = I_k - \sum \eta_j{}^I(\xi)\eta_j{}^I(\xi)'$ is a function on P with values in positive definite matrices. Moreover, for $s \in O(k)$ and $\xi \in P$,

$$R(\xi s) = s^{-1}R(\xi)s.$$

For,

$$\begin{aligned} R(\xi s) &= I_k - \sum \eta_j{}^I(\xi s)\eta_j{}^I(\xi s)' \\ &= I_k - \sum s^{-1}\eta_j{}^I(\xi)\eta_j{}^I(\xi)'(s^{-1})' \\ &= s^{-1}(I_k - \sum \eta_j{}^I(\xi)\eta_j{}^I(\xi)')s \\ &= s^{-1}R(\xi)s. \end{aligned}$$

Let $S(\xi)$ be the differentiable positive matrix-valued function on P such that

$S(\xi)^2 = R(\xi)$. It is clear from the uniqueness of the positive square root of a positive definite matrix that $S(\xi s) = s^{-1}S(\xi)s$ for $\xi \in P$, $s \in O(k)$.

Let $\psi: P \to W(n+k,k)$ be a G-morphism, the existence of which is assured by the universal bundle theorem [6, § 19]. Consider the matrix

$$\psi_1(\xi) = \begin{pmatrix} \eta_1(\xi)' \\ \vdots \\ \eta_i(\xi)' \\ \vdots \\ \eta_{(n+1)n^p}(\xi)' \\ \psi(\xi)S(\xi) \end{pmatrix}$$

where $\eta_i = \eta_j^I$ with $\lambda(I,J) = i$. Each η_i' is a $(1,k)$ matrix and $\psi(\xi)S(\xi)$ is a $(n+k,k)$ matrix so that $\psi_1(\xi)$ is a $((n+1)n^p+n+k,k)$ matrix. The map $\xi \to \psi_1(\xi)$ is a map of P into $W(N,k)$. For,

$$\begin{aligned} \psi_1'(\xi)\psi_1(\xi) &= \sum_{i=1}^{(n+1)n^p} \eta_i(\xi)\eta_i(\xi)' + S(\xi)'\psi(\xi)'\psi(\xi)S(\xi) \\ &= \sum_{(I,j)} \eta_j^I(\xi)(\eta_j^I(\xi))' + S(\xi)^2 \\ &= I_k - R(\xi) + S(\xi)^2. \end{aligned}$$

Hence $\psi_1'(\xi)\psi_1(\xi) = I_k$.

Moreover $\psi_1: P \to W(N,k)$ is a G-morphism. In fact,

$$\begin{aligned} \psi_1(\xi s) &= \begin{pmatrix} \eta_1^I(\xi s)' \\ \vdots \\ \eta_i^I(\xi s)' \\ \vdots \\ \eta_{(n+1)n^p}{}^I(\xi s)' \\ \psi(\xi s)S(\xi s) \end{pmatrix} = \begin{pmatrix} (s^{-1}\eta_1(\xi))' \\ \vdots \\ (s^{-1}\eta_{(n+1)n^p}(\xi))' \\ \psi(\xi)s\cdot s^{-1}S(\xi)s \end{pmatrix} = \begin{pmatrix} \eta_1^I(\xi s)' \\ \vdots \\ \eta_{(n+1)n^p}{}^I(\xi)' \\ \psi(\xi)S(\xi) \end{pmatrix} s \\ &= \psi_1(\xi)s. \end{aligned}$$

Finally, we construct a G-morphism Φ of P into E such that $\Phi^*\alpha_0 = \alpha$ (E and α_0 are the bundle and p-form constructed in the beginning of this section). $\Phi: P \to E$ is defined by $\pi_0 \circ \Phi = h$, $\pi_j \circ \Phi = (f_j^1 \circ q, \cdots, f_j^n \circ q)'$, $(j = 1, \cdots, n+1)$ and $\pi \circ \Phi = \psi_1$. Φ is a G-morphism since ψ_1 is so and h, $f_j^i \circ \pi$ are invariant functions. We then have

$$\begin{aligned} \Phi^*(\alpha_0) &= \Phi^*(\sum \pi_0 \tau_j^I \omega_j^I) \\ &= \sum (\pi_0 \circ \Phi)(\tau_j^I \circ \Phi)\Phi^*(\omega_j^I) \end{aligned}$$

$$= \sum h(\sigma_{\lambda(I,j)} \circ \pi \circ \Phi)\Phi^*(\omega_j^I)$$
$$= \sum h\eta_{\lambda(I,j)}\Phi^*\pi_j^*(dx_j^{i_1}\otimes\cdots\otimes dx_j^{i_p})$$

where $I = (i_1,\cdots,i_p)$ with $1 \leqq i_p \leqq n$. Hence

$$\Phi^*(\alpha_0) = \sum_{(I,j)} h\eta_j^I d(f_j^{i_1}\circ q)\otimes\cdots\otimes d(f_j^{i_p}\circ q) = \alpha.$$

5. The case of a compact group. In this section we prove Theorem 2 with G compact and ρ any k-dimensional representation of G. Since every representation of G is equivalent to an orthogonal representation, we may assume that $\rho = j\circ f$ where f is a homomorphism $G \to O(k)$ and j is the natural representation of $O(k)$ in $\boldsymbol{R}^k$. Let E_1 be a $O(k)$-bundle together with an n-universal p-form of type j (§ 4). Let F be a n-universal G-bundle. We let G act on $F\times E_1$ by $(v,e_1)g = (vg, e_1f(g))$, $v\in F$, $e_1\in E_1$, $g\in G$. This makes of $F\times E_1$ a G-bundle E. Let q_1 and q_2 be the projections of $F\times E_1$ onto F and E_1 respectively. The p-form $\alpha_0 = q_2{}^*\alpha_1$ is of type ρ since $q_2\colon E\to E_1$ is a f-morphism and α_1 of type j. We now prove that α_0 is a n-universal p-form of type ρ. In fact, let P be a G-bundle over a base of dimension $\leqq n$ and α a p-form of type ρ on P. Then there exists a $O(k)$-morphism Φ_2 of $T_f(P)$ into E_1 such that $\Phi_2{}^*(\alpha_1) = T_f(\alpha)$. (For the definition of T_f see § 2). On the other hand, P admits a G-morphism Φ_1 into F, since F is n-universal. Let $\Phi\colon P\to E$ be the map defined by $q_1\circ\Phi = \Phi_1$, $q_2\circ\Phi = \Phi_2\circ i_f$ (i_f is the canonical map $P\to T_f(P)$, see § 2). Φ is a G-morphism. For, if $\xi\in P$, $s\in G$, we have

$$\begin{aligned}\Phi(\xi s) &= (\Phi_1(\xi s), \Phi_2\circ i_f(\xi s))\\ &= (\Phi_1(\xi)s, \Phi_2(i_f(\xi)f(s)))\\ &= \Phi(\xi)s.\end{aligned}$$

Now

$$\begin{aligned}\Phi^*(\alpha_0) &= \Phi^*(q_2{}^*\alpha_1) = (q_2\circ\Phi)^*\alpha_1\\ &= (\Phi_2\circ i_f)^*\alpha_1 = i_f{}^*\Phi_2{}^*\alpha_1 = i_f{}^*T_f\alpha = \alpha.\end{aligned}$$

This completes the proof in the compact case.

6. Proof of Theorem 2. The general case. Let G be a connected Lie group and K a maximal compact subgroup of G. We denote by f the inclusion map $K\to G$. We seek to construct a n-universal p-form of type ρ, where ρ is any finite dimensional representation of G. From § 5, there exists a n-universal p-form α_1 of type $(\rho\circ f)$ on a K-bundle E_1. Then the p-form $\alpha_0 = T_f(\alpha_1)$ on the G-bundle $E = T_f(E_1)$ is n-universal for p-forms of

type ρ. In fact, let P be a G-bundle over a base of dimension $\leqq n$ and α a p-form of type ρ on P. It is well known that there exists a K-bundle P_1 such that $T_f(P_1)$ is G-isomorphic to P (reduction of structure group to K, see [6. § 12]). We identify P and $T_f(P_1)$ by such an isomorphism. Consider the form $i_f{}^*(\alpha)$ on P_1 which is of type $\rho \circ f$. Now let $\Phi_1 \colon P_1 \to E_1$ be a K-morphism such that $\Phi_1{}^*(\alpha_1) = i_f{}^*(\alpha)$. Consider the G-morphism $\Phi = T_f(\Phi_1)$ of P into E. Then we have

$$\begin{aligned}\Phi^*(\alpha) &= (T_f(\Phi_1))^*\alpha_0 = (T_f(\Phi_1))^* T_f(\alpha_1)\\ &= T_f(\Phi_1{}^*\alpha_1) = T_f(i_f{}^*\alpha)\\ &= \alpha.\end{aligned}$$

It is clear, referring to §§ 4, 5, 6, that the bundle E can be chosen to be n-classifying. This completes the proof of Theorem 2.

Remark. Theorems 1 and 2 hold even when G is a Lie group with a finite number of connected components; our proofs continue to be valid in this case.

TATA INSTITUTE OF FUNDAMENTAL RESEARCH, BOMBAY.

REFERENCES.

[1] S. S. Chern, *Topics in differential geometry*, Mimeographed notes, the Institute for Advanced Study, Princeton, 1951.
[2] J. L. Koszul, *Lectures on fibre bundles and differential geometry*, The Tata Institute of Fundamental Research, Bombay, 1960.
[3] M. S. Narasimhan and S. Ramanan, "Existence of universal connections," *American Journal of Mathematics*, vol. 83 (1961), pp. 563-572.
[4] J. Nash, "The imbedding problem for Riemannian manifolds," *Annals of Mathematics*, vol. 63 (1956), pp. 20-63.
[5] K. Nomizu, *Lie groups and differential geometry*, Publications of Mathematical Society of Japan, 1956.
[6] N. Steenrod, *The Topology of fibre bundles*, Princeton, 1951.

SYMPLECTIC GROUPS OVER LOCAL RINGS.*

By Wilhelm Klingenberg.

Let L be a local ring, i.e., a commutative ring with unit and a greatest ideal $I \neq L$. Then $L^* = L - I$ form a group under the multiplication. Let $V = V(L)$ be a $n = 2m$-dimensional vector space over L, that is, a free L-module with n generators. It turns out that there is, up to equivalence, only one non degenerate skewsymmetric bilinear form over V. The symplectic group over V, $Sp(V)$, is the group of linear automorphisms of V which leave such a form invariant.

In this paper we determine the invariant subgroups of $Sp(V)$. To simplify the proofs we restrict ourselves to the case where $\mathrm{char}(L/I) \neq 2$ and $L/I \neq \boldsymbol{F}_3$.

If L is a field then Dieudonné [2], [3] has proved that $Sp(V)$ is simple over its center, see also Artin [1]. If L is not a field then there will be other invariant subgroups, i.e., congruence subgroups with respect to the ideals J in L with $J \neq L$ and $J \neq 0$. Our principal result, Theorem 3, is that there are no other invariant subgroups in $Sp(V)$.

This conforms with out earlier results on the invariant subgroups of the linear and orthogonal groups over a local ring, cf. [4], [5]. Allowing for some exceptions which are familiar from the classical groups over a field we have therefore established *the following feature of the classical groups over a local ring: The only invariant subgroups are the congruence subgroups.*

1. General properties of the symplectic group.

1.1. Let $V = V(L)$ be a n-dimensional vector space over the local ring L. On V there shall be given a non degenerate bilinear form Φ satisfying $\Phi(X,X) = 0$ for all $X \in V$. Here non degenerate means that the homomorphism

$$d_\Phi\colon V \to V^* \tag{1.1}$$

of V into its dual V^* which is given by $d_\Phi(Y)(X) = \Phi(X,Y)$ is an isomorphism. We call V, endowed with such a form Φ, a *symplectic space* (*over* L).

* Received August 27, 1962.

Let V, V' be symplectic spaces. An isomorphism from V into V' will also be called an isometry. The group of isometries of V onto itself is called the *symplectic group over* V, $Sp(V)$. $Sp(V)$ contains, in particular, the identity map: $X \to X$ and the map: $X \to -X$. We denote these maps by 1 and -1 respectively.

Instead of $\Phi(X,Y)$ we also write XY or (XY). If $XY=0$ we call X and Y *orthogonal.*

A submodule U of the symplectic space V is called a *subspace* if (i) U is a direct summand of V and (ii) image $(d_{\Phi|U}\colon U \to U^*)$ is a direct summand of U^*. A subspace U is called *non isotropic* if kernel $(d_{\Phi|U})=0$.

Let U, U_1, U_2 be subspaces of V. We say that U is *orthogonal sum* of U_1 and U_2, notation, $U=U_1 \perp U_2$, if U is direct sum of U_1 and U_2 and if each vector of U_1 is orthogonal to each vector of U_2.

If $V=V_1 \perp V_2$ and $\sigma\colon V \to V'$ is an isometry of V into V' then σ is determined by $\sigma_1=\sigma \mid U_1$ and $\sigma_2=\sigma \mid U_2$. Therefore, we write $\sigma=\sigma_1 \perp \sigma_2$.

1.2. PROPOSITION 1.1. *Let U be a subspace of the symplectic space V. Let U^0 be the submodule of V annulated by $d_\Phi(U) \subset V^*$. Then U^0 is a subspace, called the orthogonal space of U. The following does hold*

$$(1.2) \qquad \dim U + \dim U^0 = \dim V\,;\ U^{00}=U\,;\ U \cap U^0 = \operatorname{kernel}(d_{\Phi|U})$$

Proof. Since (1.1) is an isomorphism $d_\Phi(U)$ is a direct summand of V^* and hence, U^0 is a direct summand of V with $\operatorname{codim} U^0=\dim U$. From $U^{00} \supset U$ and $\dim U=\dim U^{00}$ follows $U^{00}=U$.

Let W be a direct summand of V. Then $d_\Phi(V) \mid W = W^*$. With this one easily verifies the existence of a canonical isomorphism between $U^*/(d_\Phi(U) \mid U)$ and $U^{0*}/(d_\Phi(U^0) \mid U^0)$. This shows that

$$d_\Phi(U^0) \mid U^0 = \text{image}\ (d_{(\Phi|U^0)}\colon U^0 \to U^{0*})$$

is a direct summand of U^{0*}, i.e., that U^0 is a subspaces of V.

PROPOSITION 1.2. *Let U be a subspace of V. The following properties are equivalent:* (i) *U is non isotropic.* (ii) *U^0 is non isotropic.* (iii) $V=U \perp U^0$.

Proof. Assume (i). Then $d_{\Phi|U}\colon U \to U^*$ is an isomorphism. Therefore, to each $Y \in V$ there is a $Y_0 \in U$ with $d_\Phi(Y) \mid U = d_{\Phi|U}(Y_0) \in U^*$, i.e., $Y-Y_0 \in U^0$, i.e., we have (iii). On the other hand, if U is isotropic, it means that $U \cap U^0 \neq 0$.

1.3. An ordered pair $\{A, B\}$ of vectors in V satisfying $AB = 1$ is called *hyperbolic*. The L-module $P = \langle A, B\rangle$, spanned by a hyperbolic pair $\{A, B\}$, is a non isotropic subspace of dimension 2; it will be called *hyperbolic plane*.

PROPOSITION 1.3. *Let U be a non isotropic subspace of V. Then any $A \in U$ with $A \neq 0 \bmod I$ can be complemented in U to a hyperbolic pair.*

Proof. The assumptions imply that $d_{\Phi|U}(A) \in U^*$ is an element $\neq 0 \bmod I$. But then there is a $B \in U$ with $d_{\Phi|U}(A)(B) = BA = -1$.

THEOREM 1. *A symplectic space V is a direct sum of hyperbolic planes.*

COROLLARY. $\dim V$ *is even. Any two symplectic spaces of the same dimension are isometric.*

Proof. Since V is non isotropic, $\dim V \neq 1$. For $\dim V = 2$ apply Proposition 1.3. For $\dim V > 2$ we have from Propositions 1.3 and 1.2 that $V = P \perp P^0$, P a hyperbolic plane and $\dim P^0 = \dim V - 2$, so that we can argue by induction.

We see that V possesses a basis E_i, $1 \leqq i \leqq n = 2m$, such that $\{E_{2j-1}, E_{2j}\}$ is a hyperbolic pair, for $1 \leqq j \leqq m$. Such a basis is called *canonical*. If V and V' are symplectic spaces of the same dimension a linear map $V \to V'$ which carries a canonical basis into a canonical basis is an isometry, hence we have the corollary.

2. The congruence subgroups.

2.1. Let J be an ideal in L, $J \neq L$. Let $V = V(L)$ be a symplectic space over L. The canonical homomorphism $g_J\colon L \to L/J$ of L onto the local ring L/J determines a homomorphism

$$g_J\colon V(L) \to V(L/J) \tag{2.1}$$

of $V(L)$ onto the symplectic space $V(L/J)$ over L/J formed by the vectors $g_J X$, $X \in V(L)$, for which $\Phi(g_J X, g_J Y)$ is given by $g_J \Phi(X, Y)$. In putting $V(L/L) = 0$-module we extend (2.1) also to the case $J = L$. (2.1) determines a homomorphism

$$h_J\colon Sp(V(L)) \to Sp(V(L/J)) \tag{2.2}$$

with $h_J \sigma g_J = g_J \sigma$, for $\sigma \in Sp(V)$. Here, $Sp(V(L/L))$ shall be the unit group E.

Let J be an ideal in L. The *general congruence subgroup* mod J *of* $Sp(V)$, $GSp(V,J)$, is defined as h_J^{-1} center $Sp(V(L/J))$. The *special congruence subgroup* mod J *of* $Sp(V)$, $SSp(V,J)$, is defined as $h_J^{-1}(1)$ $=$ kernel h_J. Note: $GSp(V,L)=SSp(V,L)=Sp(V)$; $GSp(V,0)=$ center $Sp(V)$; $SSp(V,0)=E$.

Let $u\in L$. Then $o(u)$ denotes the ideal generated by u. Let $X\in V$; the *order of* X, $o(X)$, is the smallest ideal $J\subset L$ such that $g_JX=0$. Let $\sigma\in Sp(V)$; the *order of* σ, $o(\sigma)$, is the smallest ideal $J\subset L$ such that $h_J\sigma\in$ center $Sp(V(L/J))$, i.e., $\sigma\in GSp(V,J)$. Let G be a subgroup of $Sp(V)$; the *order of* G, $o(G)$, is the smallest ideal $J\subset L$ such that $G\subset GSp(V,J)$. Note: $o(X)$ is generated by the components of X with respect to any basis of V. $o(G)$ is generated by the ideals $o(\sigma)$, $\sigma\in G$.

2.2 A (*symplectic*) *transvection* is an element $\tau\in Sp(V)$ of the form

$$\tau X=X+A(AX)u \text{ with } o(A)=L \text{ and } u\in L \tag{2.3}$$

A is called *direction* of τ. One checks that τ is indeed an element of $Sp(V)$. A and u are not uniquely determined by τ: (A,u) can be replaced by any pair $(A',u')\in V\times L$ with $o(A')=L$ such that $A(AX)u=A'(A'X)u'$ for all $X\in V$. Note: $o(\tau)\subset o(u)$.

PROPOSITION 2.1. center $Sp(V)=\{1,-1\}$.

COROLLARY 1. *For* $J\neq L$ *is* $GSp(V,J)/SSp(V,J)\cong\{1,-1\}$.

COROLLARY 2. *Let* τ *be the transvection* (2.3). *Then* $o(\tau)=o(u)$.

Proof. Let $\rho\in$ center $Sp(V)$. Let τ be the transvection (2.3) with $u=1$. Then $\rho\tau\rho^{-1}=\tau$ implies $\rho A(\rho AX)=A(AX)$ for all $X\in V$, i.e., $\rho A=Aa$, where $a\in L^*$ may depend on A. Let $B\in V$ such that g_IA, g_IB are linearly independent in $g_IV=V(L/I)$. Then $\rho B=Bb$ and $\rho(A+B)=(A+B)c$ $=Aa+Bb$. Hence, $a=b=c$, i.e., $\rho X=Xa$ with the same $a\in L^*$ for all all $X\in V$. From $(XY)=(\rho X\rho Y)=(XY)a^2$ then follows $a^2=1$.

Consider now τ, (2.3). Then $h_J\tau\in\{1,-1\}$ implies $o(u)\subset J$.

2.3 PROPOSITION 2.2. *Let* B, B' *be in* V *with* $o(B)=o(B')=L$. *Then* B *can be carried into* B' *by a product of transvections of order* $\subset o(B'-B)$.

Proof 1. Assume $o(B'-B)=L$. If $(B'B)\in L^*$ then the transvection

$$\tau X=X+(B'-B)((B'-B)X)(B'B)^{-1} \tag{2.4}$$

carries B into B'. If $(B'B) \in I$ then there is a $C \in V$ with $BC \in L^*$ and $(B'C) \in L^*$. Then carry by transvections B into C and C into B'.

2. Assume $o(B'-B) = J \subset I$. There is a canonical basis E_i, $1 \leqq i \leqq n$, with $E_1 = B$. Then $B' = E_1 + \Sigma E_i e_i$ with $e_i \in J$. Put $E_1 = E_1{}^0$ and $E_1 + \Sigma_1{}^j E_i e_i = E_1{}^j$, for $j > 0$. For $j > 0$ define the transvections

$$(2.5) \qquad \begin{aligned} \tau_j X &= X + (E_2 + E_j)((E_2 + E_j)X)e_j((E_2 + E_j)E_1{}^{j-1})^{-1} \\ \tau'_j X &= X - E_2(E_2 X)e_j(E_2(E_1{}^j + E_2 e_j))^{-1}. \end{aligned}$$

Note that τ_2 is defined also for $2 \in I$, if we cancel out 2 in the numerator and denominator. Now $\tau'_j \tau_j E_1{}^{j-1} = E_1{}^j$, i.e., $\prod \tau'_j \tau_j$ carries $E_1{}^0 = B$ into $E_1{}^n = B'$. Furthermore, $o(\tau_j) = o(\tau'_j) = o(e_j) \subset J$.

Proposition 2.3. *Let $\{A, B\}$ and $\{A, B'\}$ be hyperbolic pairs. Then there exists a product of transvections of order $\subset o(B'-B)$ which leaves A invariant and carries B into B'.*

Proof 1. Assume $o(B'-B) = L$. If $B'B \in L^*$ then τ, (2.4), gives the answer. If $B'B \in I$ then $\tau'\tau$ with $\tau X = X + A(AX)$,

$$\tau' X = X - (B' - A - B)((B' - A - B)X)(1 - B'B)^{-1}$$

does the job.

2. Assume $o(B'-B) = J \subset I$. Let E_i, $1 \leqq i \leqq n$, be a canonical basis with $E_1 = -B$ and $E_2 = A$. Then $-B' = E_1 + \Sigma E_i e_i$ with $e_i \in J$, $e_1 = 0$. Put $E_1 = E_1{}^0$, $E_1 + \Sigma_1{}^j E_i e_i = E_1{}^j$ for $j > 0$. Define τ_j, τ'_j by (2.5). Then $\prod \tau'_j \tau_j$: $\{E_1{}^0, E_2\} = \{-B, A\} \to \{E_1{}^n, E_2\} = \{-B', A\}$ and $o(\prod \tau'_j \tau_j) \subset J$.

Theorem 2. (First characterization of $SSp(V, J)$) *$SSp(V, J)$ is generated by the symplectic transvections of order $\subset J$. In particular $Sp(V)$ is generated by the transvections.*

Proof. From Corollary 2 of Proposition 2.1 follows that $SSp(V, J)$ contains all transvections of order $\subset J$. We now prove that a $\sigma \in SSp(V, J)$ can be written as a product of transvections of order $\subset J$: Let E_i be a canonical basis of V, put $\sigma E_i = E'_i$. Then $o(E'_i - E_i) \subset J$. According to Proposition 2.2, E_1 can be carried into E'_1 by transvections of order $\subset J$. These transvections carry E_i, $i > 1$, into $E'_i{}'$, and we have $o(E'_i{}' - E'_i) \subset J$. Proposition 2.3 says that $E'_2{}'$ can be carried into E'_2 by transvections of order $\subset J$ such that E'_1 remains invariant. Hence, $\{E_1, E_2\}$ can be carried into $\{E'_1, E'_2\}$ by transvections of order $\subset J$. Hereby, $\langle E_1, E_2 \rangle^0$ will go into

$\langle E'_1, E'_2\rangle^0$ such that for the image $E'^{\prime\prime}_i$ of E_i, $i > 2$, $o(E'^{\prime\prime}_i - E'_i) \subset J$. Therefore, we can conclude the proof by induction on $\dim V$.

3. The structure of $Sp(V)$.

3.1 We will show that the invariant subgroups of $Sp(V(L))$ are closely related to the invariant subgroups of $SL(2, L)$, the special linear group in 2 variables over L. The invariant subgroups of this latter group have been determined in [4], for the case that $\operatorname{char}(L/I) \neq 2$ and $L/I \neq \boldsymbol{F}_3$. We will, therefore, from now on restrict ourselves to local rings L having these properties.

3.2 PROPOSITION 3.1. *Let P be a hyperbolic plane in V. Assume that $\rho \in Sp(V)$ can be written as $\rho = \rho \mid P \perp 1 \mid P^0$. Then the invariant subgroup G in $Sp(V)$, generated by ρ, contains $SSp(V, o(\rho))$.*

COROLLARY. *The invariant group, generated by a transvection τ, is $SSp(V, o(\tau))$.*

Proof. We denote by $SG(2, L, J)$ the group of linear automorphisms of P (considered as a 2-dimensional vector space) which is generated by the linear transvections of order $\subset J$, cf. [4]. In particular, $SC(2, L, L) = SL(2, L) =$ group of the special linear transformations of P. Now, since $\dim P = 2$, any linear transvection $\tau X = X + B\phi(X)$ of P can be written as a symplectic transvection $\tau X = X + A(AX)u$ with $B = Au$, $\phi(X) = (AX)$. Hence, $SC(2, L, J) = SSp(P, J)$; in particular, $SL(2, L) = Sp(P)$.

Under our assumptions, Theorem 3 in [4] implies that $SSp(P, o(\rho \mid P)) = SC(2, L, o(\rho \mid P)) \subset G$. This means, in particular, that G contains all transvections with a direction $A \in P$ and order $\subset o(\rho \mid P) = o(\rho)$. Since $Sp(V)$ operates transitively on the directions or order L it follows that G contains all transvections of order $\subset o(\rho)$. Theorem 2 gives the assertion. The corollary follows from the remark that a transvection satisfies the assumptions made for ρ.

3.2. LEMMA. *Let $\sigma \in Sp(V)$. The invariant subgroup G, generated by σ, contains $SSp(V, o(\sigma))$.*

Proof. We divide the proof into four steps.

(i) *Assume $\sigma A = Aa + Bb$ and $AB \in L^*$. Then $SSp(V, o(b)) \subset G$.* Indeed, let τ be given by (2.3) with $u = 1$. Then $\rho = \tau\sigma\tau^{-1}\sigma^{-1} \in G$. A and B span a non isotropic plane P. Since $\rho = \rho \mid P \perp 1 \mid P^0$ and $o(\rho) = o(b)$ we can apply Proposition 3.1.

(ii) *Assume $o(\sigma)=L$. Then the Lemma holds.* Indeed, there is an $A\in V$ with $o(\sigma A-A)=L$ and $\sigma A-A,A$ linearly independent mod I. If $(\sigma AA)\in L^*$ apply (i). If $(\sigma AA)\in I$ we can find a $C\in V$ with $C\sigma A=0$, $CA=1$, as can be seen from making a canonical basis with $E_1=A$. Define τ by $\tau X=X+(C-A)((C-A)X)$. Then $\rho=\tau\sigma^{-1}\tau^{-1}\sigma\in G$ and $(\rho AA)\in L^*$. Proceed as above.

(iii) *Let $o(\sigma)\subset I$. If $o(A)=L$ then $SSp(V,o(A\sigma A))\subset G$.* Indeed, let $\{A,B\}$ be a hyperbolic pair and define τ by (2.3) with $u=1$. Then $\sigma'=\sigma^{-1}\tau\sigma\tau^{-1}\in G$ and

$$\sigma'A=A+\sigma^{-1}A(A\sigma A),\ \sigma'B=B-A+\sigma^{-1}A(A(\sigma B-\sigma A)).$$

Put $-(A\sigma A)(A(\sigma B-\sigma A))^{-1}=x$. $o(x)=o(A\sigma A)$. Put $A+xB=A^*$. Then $\sigma'A^*=A^*(1-x)+Bx^2$. From (i) follows that G contains the transvection $\tau'X=X+B(BX)x^2(1-x)^{-1}$. Therefore, $\sigma^*=\tau'\sigma'\in G$. $\sigma^*A^*=A^*(1-x)$. Define τ^* by $\tau^*X=X+A^*(A^*X)$. Then $\rho^*=\sigma^{*-1}\tau^*\sigma^*\tau^{*-1}\in G$. With $1=(A^*B)=(\sigma^*A^*\sigma^*B^*)=(1-x)(A^*\sigma^*B)$ we find $\rho^*B=B+A^*y$, $y=(1-x)^{-2}-1$, i. e., $o(y)=o(x)=o(A\sigma A)$. From (i) now follows the assertion.

(iv) *Let $o(\sigma)\subset I$. Then the Lemma holds.* Indeed, let K be the ideal generated by the elements $A\sigma A$ with $o(A)=L$. Then $K\subset o(\sigma)\subset I$. We will show that $K=o(\sigma)$. Since we have from (iii) that $SSp(V,K)\subset G$ this will prove the assertion.

Let $X\in V$ and $\rho\in Sp(V)$. Then we write $\bar X$ and $\bar\rho$ for g_KX and $h_K\rho$, respectively. Assume now that there is an $A\in V$, $o(A)=L$, with $o(\bar\sigma\bar A-\bar A)=\bar J\neq(0)$, $\bar J\subset\bar I=h_KI$. Here we assume $h_{I\sigma}=1$. The case $h_{I\sigma}=-1$ can be handled in a similar way, by considering $\bar\sigma\bar A+\bar A$ instead of $\bar\sigma\bar A-\bar A$. We write $\bar\sigma\bar A-\bar A=\Sigma\bar F_\alpha\bar u_\alpha$, $1\leqq\alpha\leqq r$, where the $\bar F_\alpha$ are linearly independent mod $\bar I$ and the $\bar u_\alpha$ form a minimal system of generators for $\bar J$, cf. the proof of the Lemma in [4].

For any $C\in V$ with $o(C)=L$ we have $o(A+C)=L$ or $o(A-C)=L$, hence $(\bar A+\bar C)(\bar\sigma\bar A+\bar\sigma\bar C)=0$ or $(\bar A-\bar C)(\bar\sigma\bar A-\bar\sigma\bar C)=0$ and therefore, with $\bar A\bar\sigma\bar A=\bar C\bar\sigma\bar C=0$, $\bar C(\bar\sigma\bar A-\bar A)=(\bar\sigma\bar A-\bar A)(\bar\sigma\bar C-\bar C)/2$, We now choose a C such that $\bar C\bar F_1=1$. Then this formula yields, with

$$\bar C\bar F_\alpha=\bar a_\alpha,\bar F_\alpha(\bar\sigma\bar C-\bar C)2^{-1}=\bar w_\alpha\in\bar I:\ \bar u_1(1-\bar w_1)=\Sigma\bar u_\beta(\bar w_\beta-\bar a_\beta),$$

$\beta>1$. Since $1-\bar w_1$ is a unit in $\bar L=g_KL$ this is a contradiction to our assumptions on the $\bar u_\alpha$.

THEOREM 3. (Structure of $Sp(V)$) *The only invariant proper subgroups of the symplectic group $Sp(V)$ over $V=V(L)$ are the congruence subgroups $GSp(V,J)$ and $SSp(V,J)$, J an ideal $\neq L$.*

Here we assume $\operatorname{char}(L/I)\neq 2$ and $L/I\neq \boldsymbol{F}_3$.

COROLLARY. (Second characterization of $SSp(V,J)$)

$$SSp(V,J)=\operatorname{comm}(Sp(V),GSp(V,J))=\operatorname{comm}(Sp(V),SSp(V,J))$$

In particular, $\operatorname{comm}(Sp(V))=Sp(V)$.

Proof. Let G be an invariant subgroup of order $o(G)=J$. Then $GSp(V,J)\supset G$. According to the Lemma we have $G\supset SSp(V,o(\sigma))$ for all $\sigma\in G$. But then, since J is generated by the ideals $o(\sigma)$, $\sigma\in G$, $G\supset SSp(V,J)$. Since $GSp(V,J)/SSp(V,J)\cong\{1,-1\}$, for $J\neq L$, the theorem follows.

For the proof of the corollary we first note that the mixed commutator groups belong to $SSp(V,J)$; it remains to prove that

$$o(\operatorname{comm}(Sp(V),SSp(V,J)))=J.$$

This follows from the easily verified remark that to a given transvection τ there always exists a $\sigma\in Sp(V)$ such that $o(\tau)=o(\tau^{-1}\sigma\tau\sigma^{-1})$.

3.3 Let $P\colon Sp(V)\to Sp(V)/\operatorname{center} Sp(V)$ be the natural homomorphism. Since $GSp(V,J)/SSp(V,J)\cong\{1,-1\}$, for $J\neq L$, we have $PGSp(V,J)=PSSp(V,J)$. We denote this group by $PSp(V,J)$. From Theorem 3 follows

THEOREM 4. (Structure of $PSp(V)$) *The only invariant proper subgroups of the projective symplective group $PSp(V)$ over $V=V(L)$ are the congruence subgroups $PSp(V,J)$, J an ideal $\neq L$.*

Here we assume $\operatorname{char}(L/I)\neq 2$ and $L/I\neq \boldsymbol{F}_3$.

COROLLARY (Dieudonné [2], [3]) *If L is a field, $PSp(V(L))$ is simple.*

Appendix: On the definition of a subspace of a metric vector space.

Let V be a metric vector space (metrischer Vektorraum) over a local ring in the sense of [5], 1.1. As Mr. David Schneider kindly pointed out to me, the definition of a subspace (Unterraum) U of V has to read as follows, cf. [5], 1.1: U is a submodule of V with the following properties: (i) U is a direct summand of V, (ii) image $(d_{\Phi}|_U\colon U\to U^*)$ is a direct summand of U^*.

It then follows that kernel $(d_{\Phi}|_U : U \to U^*)$ is a direct summand of U. However, this property does not imply (ii).

All statements in [5] remain true with this new definition of a subspace, the only modifications are occurring in the proofs of Satz 1 and Satz 2 in [5]. For these modifications compare the proofs of Propositions 1.1 and 1.2 of the present paper where we did use the correct definition of a subspace already.

UNIVERSITY OF CALIFORNIA, BERKELEY.

REFERENCES.

[1] E. Artin, *Geometric algebra*, New York, 1957.
[2] J. Dieudonné, *Sur les groupes classiques*, Paris, 1948.
[3] ———, *La géométrie des groupes classiques*, Berlin, 1955.
[4] W. Klingenberg, "Lineare Gruppen über lokalen Ringen," *American Journal of Mathematics*, vol. 83 (1961), 137-153.
[5] ———, "Orthogonale Gruppen über lokalen Ringen," *American Journal of Mathematics*, vol. 83 (1961), pp. 281-320.

ON THE EXISTENCE AND UNIQUENESS OF SOLUTIONS OF THE POISSON INTERFACE PROBLEM.*

By D. P. Squier.

In various physical problems connected with heat conduction, electricity and magnetism, and fluid flow, the following problem arises: D is a simply-connected two-dimensional domain enclosed by a continuous curve C. A curve F, terminating on C, divides D into two domains D_1 and D_2; and C into arcs C_1 and C_2. It is required to find a continuous function u in $\bar{D}$ (the closure of D) such that

i) $\nabla^2 u = f_1$ in D_1, $\nabla^2 u = f_2$ in D_2

ii) $u = \bar{u}$ on C, where $\bar{u}$ is a given continuous function on C

iii) $K\partial u_1/\partial n = \partial u_2/\partial n$ on $F \cap D$, K a positive constant, $F \in C^1$, where $u = u_1$ in D_1, $u = u_2$ in D_2, and the normal derivative is in the same direction in each case.

In the following discussion, a function u satisfying these properties will be called a solution of the Poisson interface problem. When f_1 and f_2 are identically zero, u will be called a solution to the Laplace interface problem. The terms Poisson-Dirichlet or Laplace-Dirichlet will refer to the corresponding boundary value problems without interfaces. These problems belong to a class also referred to in the literature as transmission problems.

Oleinik [3] and others [5], [6], [7] have proved existence of generalized and classical solutions to much more general elliptic equations with similar boundary conditions by functional analysis and other methods. The special problem above, however, can be treated by means more elementary, viz., the iteration method of Schwarz as described in [2] together with the technique in [4]. It will be shown here that a solution to this problem exists under certain smoothness conditions on the boundary for analytic arcs F and that the solution is unique. Both proofs are based on the following maximum principle.

Theorem 1. *If u is a function continuous on $\bar{D}$ satisfying $\nabla^2 u = 0$ on $D_1 \cup D_2$, and $\partial u_2/\partial n = K\partial u_1/\partial n$ on $F \in C^2$, then the maximum (or minimum) of u cannot occur in D unless u is identically a constant.*

* Received September 11, 1962.

Proof. u cannot achieve its maximum in $D_1 \cup D_2$ unless constant. Suppose it achieves a maximum on F, say at x_0. Then, if $\boldsymbol{n}$ is the interior normal to D_1 at x_0, we have by Hopf's theorem [1] $\partial u_1/\partial n < 0$. The condition on F then requires $\partial u_2/\partial n < 0$, i.e., on the straight line normal through x_0, u is a strictly monotone function in the neighborhood of x_0, contradicting the maximality of $u(x_0)$. Thus the maximum must occur on C. A similar argument holds for the minimum.

COROLLARY. *There is at most one function continuous on $\bar{D}$ satisfying conditions* i), ii), *and* iii).

Proof. If there are two, say u and v, then $w = u - v$ satisfies the hypotheses of Theorem 1 with zero boundary values. Zero is then the maximum and minimum value of w on $\bar{D}$. Thus $w \equiv 0$ on $\bar{D}$.

LEMMA 1. *In the interface problem let F be a straight line, D be a domain symmetric with respect to that line and one for which the Laplace-Dirichlet problem is solvable. Then the Poisson interface problem is solvable for D with interface F if $f_1 \in C_\mu$ in D_1 and $f_2 \in C_\mu$ in D_1, with f_1 and f_2 bounded on $\bar{D}_1$ and $\bar{D}_2$ resp. (C_μ indicates Hölder continuity with exponent $\mu > 0$). Moreover, $u_1 \in C^1{}_\mu$* in *$D_1 \cup F$ and $u_2 \in C^1{}_\mu$ in $D_2 \cup F$. If $u^{(n)}$ is any sequence of continuous functions satisfying* i) *and* iii) *and converging uniformly on C to $\bar{u}$, then* $\lim u^{(n)} = u$ *exists and u is the solution of the Poisson interface problem with boundary values $\bar{u}$.*

Proof. Assume F is the line $y = 0$. Let w_e solve in $\bar{D}$ the problem

$$\left.\begin{aligned} \nabla^2 w_e\big|_p &= \frac{Kf_1(p) + f_2(q)}{1+K} \\ \nabla^2 w_e\big|_q &= \frac{Kf_1(p) + f_2(q)}{1+K} \end{aligned}\right\} \quad \text{in } D$$

$$\left.\begin{aligned} w_e\big|_p &= \frac{K\bar{u}_1(p) + \bar{u}_2(q)}{1+K} \\ w_e\big|_q &= \frac{K\bar{u}_1(p) + \bar{u}_2(q)}{1+K} \end{aligned}\right\} \quad \text{on } C.$$

Here p is in $D_1 \cup C_1$, and if $p = (x, y)$, then $q = (x, -y)$. Let w_0 solve in $\bar{D}$ the problem

$$\left.\begin{aligned} \nabla^2 w_0\big|_p &= \frac{f_1(p) - f_2(q)}{1+K} \\ \nabla^2 w_0\big|_q &= \frac{f_2(q) - f_1(p)}{1+K} \end{aligned}\right\} \quad \text{in } D$$

$$\left.\begin{aligned} w_0\,|_p &= \frac{\bar{u}_1(p)-\bar{u}_2(q)}{1+K} \\ w_0\,|_q &= \frac{\bar{u}_2(q)-\bar{u}_1(p)}{1+K} \end{aligned}\right\} \qquad \text{on } C.$$

These function are solution to an ordinary Poisson-Dirichlet problem for $\bar{D}$ and this problem can be solved with the help of the Green's function. Both w_e and $w_0 \in C^1{}_\mu$ in D and $\in C^2$ in $D_1 \cup D_2$ and are continuous in $\bar{D}$. $w_e(x,-y) = w_e(x,y)$ and $w_0(x,-y) = -w_0(x,y)$. Thus, $u_1 = w_e + w_0$, $u_2 = w_e + Kw_0$ satisfies i), ii), iii) and gives the solution to the Poisson interface problem. u_1 and u_2 have the regularity indicated in the statement of the lemma. For the second part of the lemma $u^{(n)}$ is decomposed into $v + w^{(n)}$ where v satisfies the Poisson interface problem with zero boundary values and $w^{(n)}$ satisfies the Laplace interface problem with given boundary values $\bar{u}^{(n)}$. With the notation as above

$$w_1{}^{(n)} = w_e{}^{(n)} + w_0{}^{(n)}$$
$$w_2{}^{(n)} = w_e{}^{(n)} + Kw_0{}^{(n)}$$

where $w_e{}^{(n)}$ and $w_0{}^{(n)}$ are harmonic in D. The boundary values of $w_e{}^{(n)}$ and $w_0{}^{(n)}$ converge uniformly and thus $w_e{}^{(n)}$ and $w_0{}^{(n)}$ tend uniformly to harmonic limits w_e and w_0; and, moreover, the partial derivatives tend to the partial derivatives of w_e and w_0 and uniformly so on compact sub-regions of D. Thus $u^{(n)} = v + w^{(n)}$ tends uniformly to a limit u; and on F, $\partial u^{(n)}/\partial y$ tends to $\partial u/\partial y$. Thus u satisfies the Poisson interface problem.

Since the Poisson interface problem is invariant under conformal mappings, the lemma above can be extended.

LEMMA 2. *In the interface problem let F be an analytic arc. If $u^{(k)}$ solves the interface problem with continuous boundary values $\bar{u}^{(k)}$, and $\bar{u}^{(k)}$ tends uniformly to $\bar{u}$, then $u^{(k)}$ tends uniformly to u and u solves the Poisson intreface problem with boundary values $\bar{u}$.*

Proof. By the maximum principle, uniform convergence on the boundary implies uniform convergence in the interior. Hence u exists and is continuous on $\bar{D}$. For the region $D_1 \cup C_1 \cup F$, $u^{(n)} = v + w^{(n)}$ where v satisfies $\nabla^2 v = f_1$ in D_1 with zero boundary values, and $w^{(n)}$ satisfies $\nabla^2 w^{(n)} = 0$ with boundary values $\bar{u}^{(n)}$. Now $w^{(n)}$ converges uniformly to a harmonic function w in D_1 and so $u = v + w$. Thus $\nabla^2 u = f_1$ in D_1; and similarly $\nabla^2 u = f_2$ in D_2. It remains to show condition iii) is satisfied. By a conformal mapping, a

segment F^* of F is mapped onto a straight line and a suitable neighborhood of F^* is mapped onto a circle having the image of F^* as diameter. The interface problem is invariant except that new f_1 and f_2 appear. By Lemma 1, $u_1 \in C^1$ in the upper half of the circle and on the interface and $u_2 \in C^1$ in the lower half. Moreover the first order partial derivatives of the sequence converge uniformly to the first order partial derivatives of u. Thus the same is true under the inverse mapping and iii) is satisfied.

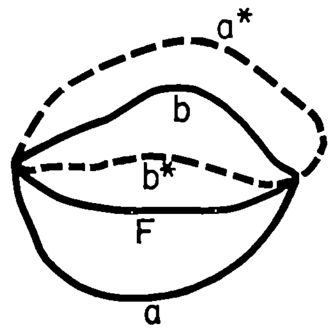
The existence of at least one solution to the interface problem is based on the following lemma, which is a slight modification of the one proved in Courant-Hilbert [2]. The proof will be only sketched.

LEMMA 3. *S_1 and S_2 are arcs enclosing a simply connected domain A. At the two points p_1 and p_2 of their intersection each arc has a one-sided tangent and the interior angle between the curves is β, $0 < \beta < 2\pi$. S_3 is any continuously differentiable arc in A intersecting S_1 and S_2 in p_1 and p_2, but meeting each arc in a non-zero angle. S_1 is a continuously differentiable arc. If v is any harmonic function in A, with $v = 0$ on S_2 and $|v| \leqq 1$ on S_1, then there is a constant q, $0 < q < 1$, depending only on the geometry, such that $|v| \leqq q < 1$ on S_3.*

Proof. If the corner has a non-zero interior angle β, then, with the corner point considered as the origin in the complex z-plane, a transformation of the form $w = z^{\pi/\beta}$ maps the corner onto a corner of angle π, i. e., removes the corner. The lemma in [2] referred to above is then applicable. (The boundary regularity in that proof actually requires that the curve carrying the "zero" values has a one-sided tangent at the point where it meets the curve carrying the "one" boundary values.)

Any domain for which the interface problem is solvable can now be extended by means of the iteration process of Schwarz. A special case is embodied in the following theorem.

THEOREM 2. *Let the interface problem be solvable in D(i. e., there is a continuous function u on $\bar{D}$ satisfying* i, ii, *and* iii) *for analytic F and with C_1 once continuously differentiable. Let D^* bounded by C^* be a domain for which $\nabla^2 v = f_1$, $v = \bar{v}$ on C^*, is solvable (i. e., v is continuous on $\bar{D}^*$). Let D^* overlap D_1 and suppose C, C^*, F have two and only two points in common, viz., the two points of intersection of F with C. If all curves at these points meet at non-zero angles, then the Poisson interface problem is solvable for the union of the domains $\mathfrak{D} = D \cup D^*$.*



Proof. Let a^* be the arc of C^* joining p_1 and p_2 which is exterior to D, and b^* the arc of C^* interior to D. Let b be the arc C_1 of C and a be the arc C_2 of C. It is supposed that b^* is once continuously differentiable. Boundary values $\bar{w}$ are prescribed on a^* and a and the Poisson interface problem is to be solved in the domain $\mathfrak{D}$ bounded by a and a^*. The problem is solved by iteration. Arbitrary continuous values are prescribed on b to complete those given on a. Then the interface problem in D is solved, the solution being $u^{(1)}$. The values of $u^{(1)}$ on b^* are used together with the given values on a^* to solve in D^* the Poisson equation. This solution is $u^{*(1)}$. The values of $u^{*(1)}$ on b are then used to solve the interface problem in D and so on. Thus a sequence $u^{(k)}$ is generated in D and a sequence $u^{*(k)}$ is generated in D^*. On b, $u^{*(k)} = u^{(k)}$ and so $u^{(k+1)} - u^{(k)} = u^{*(k+1)} - u^{*(k)}$ on b. On b^*, $u^{*(k+1)} = u^{(k)}$ and so $u^{(k)} - u^{(k-1)} = u^{*(k+1)} - u^{*(k)}$. Let M^*_k be the maximum of $| u^{*(k+1)} - u^{*(k)} |$ on b and M_k the maximum of $| u^{*(k+1)} - u^{*(k)} |$ on b^*. Let

$$v_k^* = \frac{u^{*(k+1)} - u^{*(k)}}{M_k}$$

in D^*.

Then v_k^* is harmonic in D^*, vanishes on a^* and $| v_k^* | \leqq 1$ on b^*. By Lemma 3, $| v_k^* | \leqq q < 1$ on b, and so $M_k^* \leqq q M_k$. Let $v_k = \dfrac{u^{(k)} - u^{(k-1)}}{M_{k-1}^*}$ in D. Then v_k solves a Laplace interface problem in D with $| v_k | \leqq 1$ on b and $v_k = 0$ on a. By the maximum principle, $| v_k | \leqq 1$ on b^* and so on b^*

$$| u^{*(k+1)} - u^{*(k)} | \leqq M_{k-1}^*$$

and so $M_k \leqq M_{k-1}^*$. Thus $M_k^* \leqq q M_{k-1}^*$. Consequently M_k^*, M_k tend to zero. Moreover the series

$$u^{(1)} + \sum_{k=1}^{\infty} [u^{(k+1)} - u^{(k)}] = \lim_{n \to \infty} u^{(n)}$$

is absolutely uniformly convergent (being dominated by a geometric series) and so $u = \lim u^{(n)}$ satisfies a Poisson interface problem in D. Similarly $u^* = \lim u^{*(n)}$ satisfies a Poisson problem in D^*. On b, $u^{*(k)} = u^{(k)}$ and so $u = u^*$ on b. On b^*, $u^{*(k+1)} = u^{(k)}$ and so $u^* = u$ on b^*. Thus, $u - u^*$ is zero on boundary b^* and b, and harmonic inside; and so consequently in the domain bounded by b and b^*, $u = u^*$. Thus the function w which is u^* in D^* and u in D is the solution to the Poisson interface problem for $\mathfrak{D}$.

The statement of the existence of the solution to the Poisson interface problem is contained in the following theorem:

THEOREM 3. *D is a simply-connected domain bounded by a closed Jordan curve C for which a barrier exists at each point. F is a segment of an analytic arc intersecting C in two distinct points p_1 and p_2 at non-zero interior angles. (C_1 and C_2 have one-sided tangents at p_1 and p_2.) f_1 defined in $\bar{D}_1$ and f_2 defined in $\bar{D}_2$ are bounded and in the interior of their respective domains are uniformly Hölder continuous with exponent $\mu > 0$, $\bar{u}$ is a prescribed continuous function on C. Then there exists one and only one continuous function u on $\bar{D}$ satisfying properties* i, ii, iii, *of the Poisson interface problem, i.e., solving the problem. Moreover, the first order partial derivatives are uniformly μ-Hölder continuous in compact sub-regions of $D_1 \cup F$ and $D_2 \cup F$.*

Proof. Since F is analytic (even at the endpoints) there is a univalent conformal mapping of F onto a straight line with a neighborhood of F mapped onto a neighborhood of the straight line. The Poisson interface problem is conformally invariant. The Poisson interface problem can now be solved for any region in the image plane which is symmetric with respect to the straight line as in Lemma 1. Thus, in the original plane the Poisson interface problem is solvable for a class of basic domains. Arbitrary domains as described in the theorem can be obtained by adding on to some basic domain other domains as in Theorem 2 by use of the iteration process. Thus existence is established. Uniqueness has already been proved.

Existence of solutions may be established for a domain with several non-intersecting analytic interfaces by considering it as the union of single-interface domains with overlapping boundaries and applying the Schwarz iteration process.

Existence may also be established for interface conditions

$$K_1 \partial u_1/\partial n = \partial u_2/\partial n, \qquad K_2 u_1 = u_2$$

where K_1 and K_2 are positive constants, since a maximum principle is valid here also. For suppose that $\max_F u_1$ occurs at x_0 on F. Then u_2 on F has

its maximum at x_0. But by the Hopf theorem, $u_2(x_0)$ cannot be the maximum of u_2 over D_2. Thus,

$$u_1 \leqq \max(\max \bar{u}_1, 1/K_2 \max \bar{u}_2).$$

Similarly, an examination of the minimum shows

$$u_1 \geqq \min(\min \bar{u}_1, 1/K_2 \min \bar{u}_2).$$

The same argument can be used to establish

$$| u_2 | \leqq \max(\max | \bar{u}_2 |, K_2 \max | \bar{u}_1 |).$$

The remainder of the existence proof is the same, depending on the construction of w_e and w_0 such that $u_1 = w_e + w_0$, $u_2 = K_2 w_e + K_1 w_0$ as in Lemma 1.

REFERENCES.

[1] E. Hopf, "A remark on linear elliptic differential equations of second order," *Proceedings of the American Mathematical Society*, vol. 3 (1952), pp. 791-793.

[2] R. Courant, D. Hilbert, *Methoden der Mathematischen Physik*, vol. 2, Springer, Berlin, 1937, pp. 261-269.

[3] O. A. Oleinik, "Boundary value problems for linear equations of elliptic parabolic type with discontinuous coefficients," *Izvestiya Akademii Nauk SSSR Seriya Matemetičeskaya*, vol. 25 (1961), pp. 3-20.

[4] D. P. Squier, "Regularity at the interface of solutions of elliptic equations," to appear in *Contributions to Differential Equations*.

[5] I. V. Girsanov, "The solution of certain boundary problems for parabolic and elliptic equations with discontinuous coefficients," *Doklady Akademii Nauk SSSR*, vol. 1 (1961), pp. 1373-1375. (Translation by the American Mathematical Society.)

[6] M. L. Rasulov, "The use of a contour integral in solving mixed problems for equations with discontinuous coefficients," *Doklady Akademii Nauk SSSR*, vol. 1 (1961), pp. 187-189. (Translation by the American Mathematical Society.)

[7] Sergio Campanato, "Osservazioni sul problema di trasmissione per equazioni differenziali lineari del seconde ordine," *Richerche di Matematica* (Naples), vol. 9 (1960), pp. 43-57.

FUNCTIONS WITH A MEAN VALUE PROPERTY II.*

By Leopold Flatto.

1. Introduction. Let $f(x)$ be a real valued continuous function defined in an n-dimensional region R and satisfying the mean value property (m. v. p.)

$$f(x) = \int_Y f(x+ty)\,d\mu(y) \quad \text{for } x \in R,\ 0 < t < \epsilon_x. \tag{1.1}$$

Here x and y are abbreviations for the vectors $(x_1, \cdots, x_n)$, $(y_1, \cdots, y_n)$. μ denotes a positive Borel measure of compact support, Y being the support of μ and $\mu(Y) = 1$. It is assumed that μ is not contained in an $n-1$ dimensional hyperplane.

We have shown in an earlier paper [8] that the m. v. p. (1.1) is equivalent to having $f(x)$ satisfy an infinite number of homogeneous partial differential equations with constant coefficients. Friedman and Littman [9] have subsequently obtained another system of equations equivalent to the m. v. p. This system proves to be more convenient to work for the problems which we will discuss. The equations in question are

$$P_j(\partial/\partial x)f = 0 \qquad (1 \leqq j < \infty) \tag{1.2}$$

where $P_j(x) = \int_Y (x\cdot y)^j d\mu(y)$, $x\cdot y = x_1y_1 + \cdots + x_ny_n$,

$$\partial/\partial x = (\partial/\partial x_1, \cdots, \partial/\partial x_n).$$

It follows furthermore from the work in [8] that the solution space F of (1.1) is finite dimensional if and only if the system of algebraic equations

$$P_j(x) = 0 \qquad (1 \leqq j < \infty) \tag{1.3}$$

has $0 = (0, \cdots, 0)$ as its only common complex root. (1.3) is equivalent to having $\sum_{j=0}^{\infty} P_j(x)z^j/j! = 1$ for all complex z. Since

$$\sum_{j=0}^{\infty} P_j(x)z^j/j! = \int_Y \sum_{j=0}^{\infty} (zx\cdot y)^j d\mu(y) = \int_Y e^{zx\cdot y}\,d\mu(y)$$

(1.3) is equivalent to

* Received September 10, 1962.

$$\int_Y e^{zx\cdot y}\, d\mu(y) = 1 \tag{1.4}$$

for all complex z. We will use formulation (1.4) in Section 2.

In this paper, we study system (1.2) and condition (1.3) in greater detail; we will occasionally quote results from [8] without proof. In Section 2 we discuss the "size" of the solution space F. One effective way of doing this is by counting for each integer $m \geqq 0$ the number of linearly independent homogeneous polynomials of degree m which satisfy (1.2). Using the work of Fischer on systems of partial differential equations [7] and certain results by Hilbert on polynomial ideals [11], we show that n distinct cases present themselves. We then proceed to illustrate these cases by various examples.

Condition (1.3) or its equivalent (1.4) seems to be difficult to verify even for Euclidean measures of familiar figures. In this direction we achieve greater success if we make some simplifying assumptions concerning the measure μ. In Sections 3 and 4 we consider respectively the following types of measures: a) μ is a discrete measure, i. e. $\mu(p_i) \neq 0$ on a finite set of points $p_i (1 \leqq i \leqq N)$, $\mu = 0$ elsewhere; b) μ is invariant under a finite irreducible group G of orthogonal transformations generated by reflections. I. e. $\mu(TY) = \mu(Y)$ for any Borel set Y where $TY = \{x \mid x = Ty, y \in Y, T \in G\}$. We refer to G as a finite irreducible o. g. g. r.

We consider in fact the following generalization of the m. v. p.

$$\int_Y f(x + ty)\, d\nu(y) = 0 \text{ for } x \in R,\ 0 < t < \epsilon_x,\ \nu(Y) = 0. \tag{1.5}$$

where ν is a discrete signed Borel measure with compact support Y. We obtain a geometric criterion on ν which guarantees that the solution space to (1.5) be finite dimensional.

For case b) we shall require some results of Coxeter [6] and Steinberg [15] on the invariants of a finite irreducible o. g. g. r. It turns out that the solution space F is finite dimensional and can be completely specified provided μ satisfies a certain moment condition. This condition is in general difficult to verify and in Section 5 we specialize μ to be a measure distributed over a regular polyhedron of fixed orientation. We thus generalize to n dimensions the following problem first studied by Walsh [17] in two dimensions and then by Beckenbach and Reade [1], [2] in three dimensions. To determine the class of functions $f(x)$ continuous in a region R such that the value of $f(x)$ at the center of a regular polyhedron in R similar to a given one be equal to the average of the values of $f(x)$ at the vertices.

The work in [2] is incomplete for the dodecahedron and icosahedron.

Using the method of Section 4, we complete this gap in Section 6. We also determine in Section 5 for the measures considered there the solution space of

$$(1.6)\qquad f(x)=\int_Y f(x+ty)\,d\mu(Ty)\ \text{ for } x\in R,\ 0<t<\epsilon_x$$

where T is arbitrary rotation (by a rotation we mean an orthogonal transformation of determinant $+1$). We will say that these functions have the rotated mean value property (r. m. v. p.).

2. The "size" of the solution space. We employ the following notation. The reader is referred to [12] for the terminology which we use here. $(P_1,\cdots,P_r)$ stands for the ideal generated by the polynomials $P_1,\cdots,P_r$. $P\mid Q$ means that the polynomial P divides the polynomial Q while $P\nmid Q$ means that P does not divide Q. With any ideal $\mathfrak{A}$ we associate its manifold of zeros $M_{\mathfrak{A}}$ and $\dim\mathfrak{A}$ stands for the dimension of $\mathfrak{A}$ as well as of $M_{\mathfrak{A}}$. If $\mathfrak{A}$ is a homogeneous ideal, then hom. dim. $\mathfrak{A}$ stands for the homogeneous dimension of $\mathfrak{A}$ ($=\dim.\mathfrak{A}-1$). Hom. dim. $\mathfrak{A}=-1$ if and only if $M_{\mathfrak{A}}$ consist of $0=(0,\cdots,0)$ only; we say in this case that $\mathfrak{A}$ is trivial. ([12], p. 103)

We will need the following two theorems on polynomial ideals due to Fischer [7] and Hilbert ([12], pp. 161-165) respectively.

THEOREM 2.1. *Let $\mathfrak{A}$ be a homogeneous ideal of polynomials in the variables $x_1,\cdots,x_n$. Let F be the space of polynomials $p(x)$ satisfying the system of partial differential equations $a(\partial/\partial x)\cdot p=0$, $a\in\mathfrak{A}$. Then the vector space V_m of homogeneous polynomials in n variables of degree m is the direct sum of the vector spaces A_m and F_m where $A_m=\mathfrak{A}\cap V_m$, $F_m=F\cap V_m$. Symbolically $V_m=A_m\oplus F_m$.*

THEOREM 2.2. *Let $H(m;\mathfrak{A})=\dim V_m-\dim A_m$ where* $\dim V_m$, $\dim A_m$ *stand respectively for the linear dimensions of V_m and A_m; let $d=$* hom. dim. $\mathfrak{A}$. *Then for sufficiently large m, $H(m;\mathfrak{A})=h_0m^d+h_1m^{d-1}+\cdots+h_d$ $(h>0)$.*

If $d=-1$, then the above statement is to be interpreted as saying that $\mathfrak{A}$ contains all homogeneous polynomial of sufficiently high degree. If $H=(\phi_1,\cdots,\phi_r)$ where $d=n-r-1$ $(r<n)$ and ϕ_i is a homogeneous polynomial of degree d_i $(1\leqq i\leqq r)$, then $h_0=\dfrac{\alpha_1\alpha_2\cdots\alpha_r}{d!}$ for $m>\alpha_1+\cdots+\alpha_r-n$.

The last statement in Theorem (2.2) implies the following:

COROLLARY. *Suppose $\mathfrak{A}=(\phi_1,-,\phi_n)$ is a trivial ideal, the ϕ_i's being homogeneous and $\alpha_i=\deg\phi_i$ $(1\leqq i\leqq n)$. Then the linear dimension of*

the residue class ring $R/\mathfrak{A}$ is $\alpha_1\alpha_2\cdots\alpha_n$, where R denote the ring of polynomials in $x_1, x_2, \cdots, x_n$. Furthermore $\mathfrak{A}$ contains all homogeneous polynomials of $\deg > \sum_{i=1}^{n} \alpha_i - n$.

Proof. Let $\mathfrak{B} = (\phi_1, \cdots, \phi_{n-1})$. By Theorem (2.2)

$$H(m, \mathfrak{B}) = \alpha_1\alpha_2\cdots\alpha_{n-1} \text{ for } m > N = \sum_{i=1}^{n-1} \alpha_i - n.$$

Now $H(m; \mathfrak{A}) = H(m; \mathfrak{B}) - H(m - \alpha_n; \mathfrak{B})$ ([11], p. 158) (By definition $H(m; \mathfrak{B}) = 0$ for $m < 0$). In particular $H(m; \mathfrak{A}) = 0$ for $m > N + \alpha_n = \sum_{i=1}^{n} \alpha_i - n$. Hence

$$\begin{aligned}
\dim R/H &= \sum_{m=0}^{N+\alpha_n} \dim(m; \mathfrak{A}) \\
&= \sum_{m=0}^{N+\alpha_n} \dim(m; \mathfrak{B}) - \sum_{m=0}^{N+\alpha_n} \dim(m - \alpha_n; \mathfrak{B}) \\
&= \sum_{m=0}^{N+\alpha_n} \dim(m; \mathfrak{B}) - \sum_{m=0}^{N} \dim(m; \mathfrak{B}) \\
&= \sum_{m=N+1}^{N+\alpha_n} \dim(m; \mathfrak{B}) = \alpha_1\alpha_2\cdots\alpha_m.
\end{aligned}$$

Combining Theorems (2.1) and (2.2) we obtain the following result.

THEOREM 2.3. *Let $g(m)$ = dimension of the space of homogeneous polynomials which satisfy the m.v.p. (1.1). Let $d = \text{hom.dim.}\,\mathfrak{P}$ where $\mathfrak{P}$ is the homogeneous ideal generated by the polynomials P_j $(1 \leqq j \leqq \infty)$. Then for sufficiently large m, $g(m)$ is a polynomial of degree d.*

Every solution (1.1) or its equivalent (1.2) can be expanded into a series of homogeneous polynomials which themselves satisfy (1.1), this being so as $P_2(\partial/\partial x) \cdot f = 0$ is an elliptic equation. We thus see again that the solution space F of (1.1) is finite dimensional if and only if $d = -1$.

We note that in previous work on the m.v.p. problem ([4], 8.9) the emphasis was on whether the solution of (1.1) form a finite or infinite dimensional space. We now see that the case of an infinite dimensional solution space should be divided into $n-1$ subcases so that in general n possibilities present themselves as d may assume any of the n values $-1 \leqq d \leqq n-2$.

The following theorems concern themselves with the calculation of hom. dim. $\mathfrak{P}$ for certain types of measures. We introduce the following notation. $G = G_1 \times G_2 \times \cdots \times G_r$ is the group of orthogonal transformations with matrix representation

$$\begin{bmatrix} T_1 & & & & & & \\ & T_2 & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & T_k & \\ & & & & & & E \end{bmatrix}$$

where the elements not shown are zeros, and where T_j $(1 \leqq j \leqq k)$ is an arbitrary orthogonal matrix of degree d_j. E is the identity matrix of degree d_{k+1}, $\sum_{j=1}^{k+1} d_j = n$. G is thus the direct product of the orthogonal groups G_j $(1 \leqq j \leqq k)$; G_j acts on the variables x_s, $a_{j-1} < s \leqq a_j$, where

$$a_j = d_1 + d_2 + \cdots + d_j, \quad S = S_1(r_1) \times S_2(r_2) \times - \times S_k(r_k)$$

denotes the direct product of the spheres $S_i(r_i)$: $\sum_{s=a_{i-1}+1}^{a_i} x_s^2 \leqq r_i^2$ $(1 \leqq i \leqq k)$ where $d_1 + d_2 + \cdots + d_k = n$.

We can now state the following result.

THEOREM 2.4. a) *Let μ be invariant under the group G. Then* hom. dim. $\mathfrak{P} \geqq \sum_{i=1}^{k} d_i - k - 1$.

b) *Let μ denote the Euclidean measure of S divided by the total measure of S (normalized Euclidean measure). Then* hom. dom. $\mathfrak{P} = n - k - 1$. *Furthermore if $f(x)$ satisfies the m.v.p.* 11.1), *then $f(x)$ is polyharmonic in the variables x_s $(a_{j-1} < s \leqq a_j,\ 1 \leqq j \leqq k)$.*

Proof. a) let T be an arbitrary rotation of the group G. Then

$$P_m(Tx) = \int_Y (Tx, y)^m, d\mu(y) = \int_Y (x \cdot T^{-1}y)^m \, d\mu(y)$$

$$= \int_Y (x \cdot y)^m \, d\mu(Ty) = \int_Y (x \cdot y)^m \, d\mu(y) = P_m(x) \tag{2.1}$$

$$(1 \leqq m < \infty).$$

Thus $P_m(x)$ is an invariant of the group G. It follows ([18], p. 53) that P_m is a polynomial in $\rho_1^2, \cdots, \rho_k^2, x_{a_{k+1}}, \cdots, x_n$ where $\rho_j^2 = \sum_{s=a_j+1}^{a_j} x_s^2$, $1 \leqq j \leqq k$; say

$$P_m(x) = Q_m(\rho_1^2, \cdots, \rho_k^2; x_{a_k+1}, \cdots, x_n) \tag{2.2}$$

Hence $P_m(x) = 0$ on the manifold M: $\rho_j = 0$ $(1 \leqq j \leqq k)$, $x_s = 0$ $(a_k < s \leqq n)$. Since hom. dim. $M = \sum_{j=1}^{k} d_j - k - 1$, it follows that hom. dim. $\mathfrak{P} \geqq \sum_{j=1}^{k} d_j - k - 1$.

b) Suppose that $P_m(x) = 0$ $(1 \leqq m < \infty)$. As shown in Section 1, this is equivalent to (2.3) $\int_S e^{zx\cdot y} d\mu(y) = 1$ for all complex z. But $\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_k$ where μ_i is the normalized Euclidean measure of the sphere $\sum_{s=a_{i-1}+1}^{a_i} x_s^2 = 1$. Hence

$$\int_S e^{zx\cdot y} d\mu(y) = \prod_{j=1}^{k} \int_S e^{zX_j\cdot Y_j} d\mu_s(Y_j) \tag{2.4}$$

where X_j, Y_j denote respectively the vectors $(x_{a_{j-1}+1}, \cdots, x_{a_j})$, $(y_{a_{j-1}+1}, \cdots, y_{a_j})$. A direct computation yields

$$\int_{S_j(r_j)} e^{zX_j\cdot Y_j} d\mu_j(Y_j) = B_{q_j}(\sqrt{-1} d_j\rho_j z) \tag{2.5}$$

where $q_j = d_j/2$ and

$$B_q(z) = \frac{2^q\Gamma(q+1)J_q(z)}{z^q}$$

$J_q(z)$ denoting the Bessel function of order q. Hence

$$\int_s e^{zx\cdot y} d\mu(y) = \prod_{j=1}^{k} B_{q_j}(\sqrt{-1}\, r_j\rho_j z) \tag{2.6}$$

Thus (2.3) becomes

$$\prod_{j=1}^{k} B_{q_j}(\sqrt{-1}\, r_j\rho_j z) = 1) \tag{2.7}$$

for all z. Since $B_{q_j}(\rho_j z)$ has zeros for $\rho_j \neq 0$, it follows that (2.7) holds if and only if $\rho_j = 0$ $(1 \leqq j \leqq k)$. Thus the manifold $M_{\mathfrak{p}}$ is given by $\rho_1^2 = \rho_2^2 = \cdots = \rho_k^2 = 0$ so that hom. dim. $\mathfrak{p} = n - k - 1$. Since $\rho_j^2 = 0$ on $M_{\mathfrak{p}}$ for $1 \leqq j \leqq k$, we conclude by Hilbert's Nullstellensatz (see [12], p. 47) that there exists a positive integer N so that $\rho_j^{2N} \in B$ $(1 \leqq j \leqq k)$. This means that $\rho_j^{2N}(x)$ $(1 \leqq j \leqq k)$ is a linear combination with polynomial coefficients of the $P_m(x)$'s. Substituting in this identity $\partial/\partial x$ for x, we conclude that $\rho_j^{2N}(\partial/\partial x)\cdot f = 0$ $(1 \leqq j \leqq k)$ so that f is polyharmonic in the variables $x_{a_{j-1}+1}, \cdots, x_{a_j}$ $(1 \leqq j \leqq k)$.

We notice in particular that if $d_j = -1$ $(1 \leqq j \leqq k)$, then hom. dim. $\mathfrak{p} = -1$ so that solution space is finite dimensional when S is a parallelopiped and μ is the normalized Euclidean volume. We will return to this problem in Section 5 were we will obtain some additional results.

As a final illustration, we characterize the solution space of (1.1) when μ is invariant under the group $G = G_1 \times G_2$ where G_1, G_2 are the orthogonal groups acting respectively on the sets of variables $(x_1, \cdots, x_{n-1}), x_n$. We set

$$A_{1,0} = \int_Y y_1 \; d\mu(y), \quad A_{11} = \int_Y y_n^2 \, d\mu(y), \quad A_{20} = \int_Y y_1^4 \, d\mu(y),$$
$$A_{21} = c \int_Y y_1^2 y_n^2 \, d\mu(y), \quad A_{22} = \int_Y y_n^4 \, d\mu(y)$$

and

$$c(\mu) = A_{20}A_{11}^2 - A_{21}A_{11}A_{10} + A_{22}A_{10}^2,$$
$$\omega = \omega(\mu) = \int_Y y_1^2 \, d\mu(y) / \int_Y y_n^2 \, d\mu(y).$$

We assume that the region R is x_n-convex i; i.e. any line joining two points of R and parallel to the x_n-axis is contained in R. This assumption is made in order to avoid any multiple valuedness for the functions u_0, u_1 introduced in the next theorem. We state the following result.

THEOREM 2.5. *Let μ be invariant under the group $G_1 \times G_2$ and suppose that $c(\mu) \neq 0$. Then the solution space of* (1.1) *is given by* $f(x) = u_0 + u_1 x_n - \omega \dfrac{\nabla u_0}{2!} x_n^2 - \omega \dfrac{\nabla u_1}{3!} x_n^3$ *where* $u_i = u_i(x_1, \cdots, x_{n-1})$ $(i = 1, 2)$ *are arbitrary biharmonic functions and* $\nabla = \sum_{i=1}^{n-1} \partial^2/\partial x_i^2$. *If μ is the normalized Euclidean volume of the cylinder* $C: \sum_{i=1}^{n-1} x_i^2 \leqq r^2$, $|x_n| \leqq h$ $(r > 0,$ $h > 0)$ *then* $c(\mu) \neq 0$.

Proof. μ remains invariant under the transformation $y_i' = -y_i$, $y_j' = y_j$ $(j \neq i)$. Hence $\int_Y y_1^{a_1} \cdots y_n^{a_n} \, d\mu(y) = 0$ if any of the a_i's is a positive odd integer. It follows that $P_{2j-1}(x) = 0$ $(1 \leqq j < \infty)$. A direct computation yields

$$P_{2j}(x) = \sum_{k=0}^{j} A_{jk} \rho^{2(j-k)} x_n^{2k} \tag{2.8}$$

where $\rho^2 = x_1^2 + \cdots + x_{n-1}^2$. Eliminating ρ^2 between P_2 and P_4 we obtain $P_4(x) \equiv \dfrac{c(\mu)}{A_{10}^2} x_n^4 \pmod{P_2(x)}$. Suppose now that $c(\mu) \neq 0$; then (P_2, P_4)

$=(P_2, x_n^4)$. It follows from (2.8) that $P_{2j}\in(P_2, x_n^4)$ $(1\leqq j<\infty)$. Thus $P=(P_2, x_n^4)$ so that system (1.2) is equivalent to

$$
(2.9)\qquad \begin{cases} \left(A_{10}\nabla + A_{11}\dfrac{\partial^2}{\partial x_n{}^2}\right)\cdot f = 0 \\[2ex] \dfrac{\partial^4}{\partial x_n{}^4} f = 0. \end{cases}
$$

We conclude from the second equation in (2.9) that

$$
(2.10)\qquad f(x) = u_0 + u_1 x_n + u_2 x_n{}^2 + u_3 x_n{}^3
$$

where $u_i = u_i(x_1,\cdots,x_{n-1})$ $(0\leqq i\leqq 3)$.

Substituting (2.10) into the first equation of (2.9) we obtain

$$
(2.11)\qquad \begin{aligned}(A_{10}\nabla u_0 + 2A_{11}u_2) + (A_{10}\nabla u_1 + 3\cdot 2A_{11}u_3)x_n \\ + (A_{10}\nabla u_2)x_n{}^2 + (A_{10}\nabla u_3)x_n{}^3 = 0.\end{aligned}
$$

Setting the coefficients in (2.11) equal to zero, we obtain the desired result

$$
(2.12)\qquad f(x) = u_0 + u_1 x_n - \omega\frac{\nabla u_0}{2!}x_n{}^2 - \omega\frac{\nabla u_1}{3!}x_n{}^3
$$

where $\nabla^2 u_0 = \nabla^2 u_1 = 0$.

We must still show that $c(\mu)\neq 0$ when μ is the normalized Euclidean measure of the cylinder $C\cdot\sum_{i=1}^{n-1} x_i^2 \leqq .r^2$, $|x_n|\leqq h$. In this case the A_{ij}'s are all Dirichlet integrals and can be computed explicitly ([19], p. 252). We list the computed values

$$
A_{10}=\frac{r^2}{n+1},\ A_{11}=\frac{h^2}{3},\ A_{20}=\frac{3r^4}{(n+1)(n+3)},\ A_{21}=\frac{2r^2h^2}{n+1},\ A_{22}=\frac{h^4}{5}.
$$

Substituting these values into the expression for $c(\mu)$, we obtain

$$
c(\mu) = -\frac{2(n+8)}{15(n+1)^2(n+3)}\, r^4h^4 \neq 0.
$$

3. The case of discrete measure. We proceed to study the solution space of

$$
(3.1)\qquad \int_Y f(x+ty)\,d\nu(y) = 0 \text{ for } x\in R,\ 0<t<\epsilon_x
$$

where $\nu(Y)=0$, Y being the support of the signed Borel measure ν. The following theorem shows the connection between (3.1) and the system of equations.

$$P_j(\partial/\partial x)\cdot f=0 \qquad (1\leqq j<\infty) \tag{3.2}$$

where $P_j(x)=\int_Y (x\cdot y)^j d\nu(y)$.

THEOREM 3.1. *Let $f(x)$ be an analytic solution of* (3.1). *Then $f(x)$ satisfies* (3.2). *Conversely if $f(x)$ satisfies* (3.2), *then there exists a sequence of C^∞ functions $f_n(x)$ $(n=1,2,\cdots)$ which satisfy* (3.2) *such that the f_n's converge uniformly to f on every compact subset of R.*

We omit the proof to the above theorem as it is identical with the proof to Theorem (2.1) in (8). We just mention that we may define $f_n(x)$ as $f_n(x)=\int\delta_n(x-y)f(y)\,dy$ where

$$\delta_n(x)=\begin{cases} c_n\exp\left(\dfrac{-1}{|x|^2-1/n^2}\right), & |x|=\sqrt{x_1^2+\cdots+x_n^2}\leqq 1/n \\ 0 & , \ |x|>1/n \end{cases}$$

c_n is a positive constant so chosen that $\int_{|x|\leqq 1/n}\delta_n(x)\,dx=1$.

We have the following

COROLLARY. *The solution space to* (3.1) *is finite dimensional if and only if the system of algebraic equations* (3.2) $P_j(x)=0$ *has* $0=(0,\cdots,0)$ *as its only common complex root.*

Proof. If $P_j(z)=0$ $(1\leqq j<\infty)$ for some $z\neq 0$, then $Re(z\cdot x)^m$ $(1\leqq m<\infty)$ is an analytic solution of (3.2) and thus a solution of (3.1). Suppose 0 is the only common root of $P_j(x)=0$ $(1\leqq j<\infty)$. As in (8), any C^∞ solution of (3.2) is a polynomial of degree $<K$, K being a positive integer determined by system (3.2). By Theorem (3.1), $f(x)$ can be approximated uniformly on compact subsets of R by polynomials of degree $<K$. It then follows from a standard result that $f(x)$ is a polynomial of degree $<K$. The same reasoning as in Section 1 shows that (3.2) is equivalent to (3.3) $\int_Y e^{zx\cdot y}\,d\nu(y)=0$ for all complex z.

Let us suppose now that ν is discrete and concentrated on N points $y_1,\cdots,y_N$ where $\nu(y_j)=\nu_j$ $(1\leqq j\leqq N)$, $\nu=0$ elsewhere. (3.1) then becomes

$$\sum_{j=1}^{N}\nu_j f(x+ty_j)=0 \text{ for } x\in R,\ 0<t<\epsilon_x \tag{3.4}$$

The following theorem includes a result found in (9).

THEOREM 3.2. *The necessary and sufficient condition that the solution space of* (3.4) *be infinite dimensional is that there exist an non-zero real vector u such that $\nu(P_c)=0$ for all real c. Here*

$$P_c=\{x \mid u_1x_1+\cdots+u_nx_n=c\}.$$

Proof. Suppose that there exists a non-zero real vector u such that $\nu(P_c)=0$ for all real c. Let $u\cdot y_j$ $(1\leqq j\leqq N)$ take on the distinct values $a_1,\cdots,a_k$. Then

$$\sum_{j=1}^{N}\nu_j[u\cdot(x+ty_j)]^m=\sum_{j=1}^{K}(u\cdot x+ta_j)^m\nu(Pa_j)=0, \qquad (1\leqq m<\infty)$$

so that $(u\cdot x)^m$ is a solution of (3.4). The solution space of (3.4) is thus infinite dimensional.

Conversely, let the solution space of (3.4) be infinite dimensional. By the corollary to Theorem (3.1), it follows that $\sum_{j=1}^{N}\nu_je^{(x\cdot y_j)z}=0$ for all complex z and some $x\neq 0$. We may assume that $u=Re\,x\neq 0$. Let $u\cdot y_j$ $(1\leqq j\leqq N)$ take on the distinct values $a_1,\cdots,a_k$. Then

$$\sum_{j=1}^{N}\nu_je^{(x\cdot y_j)z}=\sum_{j=1}^{K}\nu(Pa_j)Q_j(z)$$

where $Q_j(z)=\sum_{r=1}^{N}e^{(a_j+ib_jr)z}$ and where the b_j's are real and $n_1+\cdots+n_k=n$. Since the Q_j's $(1\leqq j\leqq K)$ form a linearly independent set of functions, $\nu(P_{a_j})=0$, $(1\leqq j\leqq K)$. $\nu(P_c)=0$ for $c\neq a_j$ $(1\leqq j\leqq K)$, for in that case P_c contains no y_j. Hence $\nu(P_c)=0$ for all real c.

As an example of an infinite dimensional solution space to (3.4), consider the m.v.p.

$$f(x_1+ty_1,x_2+ty_2)-f(x_1+ty_1,x_2)-f(x_1,x_2+ty_2)+f(x_1,x_2)=0$$

where x_1, x_2, t take on all real values. It is well known that the solution space is given by the functions $f(x_1,x_2)=g(x_1)+h(x_2)$, g and h being arbitrary continuous functions. The measure ν is concentrated in the two hyperplanes P_1 $(x_1=0)$ and P_2 $(x_1=y_1)$.

It is shown in (10) that if $f(x)$ satisfies

$$(3.5) \qquad f(x)=\sum_{j=1}^{N}\mu_jf(x+ty_j),\ \mu_j>0\ ((1\leqq j\leqq N),\ \sum_{j=1}^{N}\mu_j=1$$

and if the y_j's are not contained in a hyperplane, then $f(x)$ is a polynomial of degree $<N(N-1)/2$. It will be shown in Section 5 that this bound is sharp and is attained when the y_j's for N vertices of an $(N-1)$ dimensional

regular tetrahedron, $\mu(y_j)=1/N$ $(1\leqq j\leqq N)$. When the N points y_j $(1\leqq j\leqq N)$ are situated in two dimensions we get the following stronger result.

THEOREM 3.2. *Let*

$$f(x)=\sum_{j=1}^{N}\mu_j f(x+ty_j),\ \mu_j>0\ (1\leqq j\leqq N),\ \sum_{j=1}^{N}\mu_j=1. \tag{3.6}$$

Here $x=(x_1,x_2)$, $y_j=(y_{j1},y_{j2})$; it is assumed that μ is not contained in a line through the origin. Then $f(x)$ is a polynomial of degree $\leqq N$. If $\mu_j=1/N$ $(1\leqq j\leqq N)$, then there exists a polynomial solution of degree N to (3.6) if and only if there exist a non-singular matrix A such that

$$Ay_j=v_j \qquad (1\leqq j\leqq N) \tag{3.7}$$

where the v_j's are the vertices of a regular n-sided polygon whose center is at the origin.

Before establishing this result, we prove the following.

LEMMA. *The solution to the system of equations*

$$\sum_{j=1}^{N} z_j^k=0 \qquad (1\leqq k\leqq N-1) \tag{3.8}$$

is given by (3.9) $z_j=Ce^{2\pi i\sigma(j)/N}$ $(1\leqq j\leqq N-1)$ *where* C *is arbitrary and* $\sigma(j)$ *is a permutation of* $1,\cdots,N$.

Proof. Using the Newton identities ([16], p. 81 exercises), we establish readily that the equations (3.8) are equivalent to the equations

$$\sigma_k(z)=0 \qquad (1\leqq k\leqq N-1) \tag{3.10}$$

where

$$\sigma_k(z)=\sum_{1\leqq i_1<i_2<\cdots<i_k\leqq N} z_{i_1},\cdots,z_{i_k}$$

Now

$$\prod_{i=1}^{N}(z-z_i)=z^N-\sigma_1 z^{N-1}+\sigma_2 z^{N-2}-\cdots+(-1)^N\sigma_N.$$

If $\sigma_k=0$ $(1\leqq k\leqq N-1)$, then

$$\prod_{i=1}^{N}(z-z_i)=z^N+(-1)^N\sigma_N \tag{3.11}$$

so that (3.9) holds with $C^N=(-1)^{N+1}\sigma_N$. Conversely (3.9) is clearly a solution of (3.8).

We now proceed with the proof of Theorem (3.2).

Proof. We may assume that $P_1(x) = 0$ for otherwise the solutions to (3.6) are linear functions [4] and there is nothing to prove. $P_2(x)$ is a positive definite form, say $P_2(x) = Ax_1^2 + Bx_1x_2 + Cx_2^2$. Suppose that $P_2 \nmid P_M$ for some $M > 2$. We claim that $p(x) \in (P_2, P_M)$ whenever p is a homogeneous polynomial of degree $s > M$. For choose a linear transformation $x = Tx'$ such that $Q_2(x') = P_2(Tx') = x_1'^2 + x_2'^2$. Since $x_1'^2 \equiv x_2'^2 \pmod{Q_2}$, $Q_M(x') = P_M(Tx') \equiv Dx_1'^M + Ex_1'^{M-1}x_2'$; since $Q_2 \nmid Q_M$ we have $D^2 + E^2 \neq 0$. Let $\mathfrak{A} = (Q_2, Q_M)$. We have

$$(3.12) \qquad x_1'^{M-1}(Dx_1' + Ex_2') \equiv 0 \pmod{\mathfrak{A}}.$$

Multiplying the congruence (3.12) by $Dx_1' - Ex_2'$ we have

$$(3.13) \qquad x_1'^{M-1}(D^2x_1'^2 - E^2x_2'^2) \equiv (D^2 + E^2)x_1'^{M+1} \equiv 0 \pmod{\mathfrak{A}}$$

so that $x_1'^{M+1} \in \mathfrak{A}$.

If $E = 0$, then we conclude from (3.12) that $x_1'^M \in \mathfrak{A}$; hence $x_1'^M x_2' \in \mathfrak{A}$. If $E \neq 0$, then we multiply (3.12) by x_1' and conclude from (3.13) that $x_1'^M x_2' \in \mathfrak{A}$. Since either $x_1'^{j_1}x_2'^{j_2} \equiv \pm x_1'^{j_1+j_2} \pmod{\mathfrak{A}}$ or

$$x_1'^{j_1}x_2'^{j_2} \equiv \pm x_1'^{j_1+j_2-1}x_2' \pmod{\mathfrak{A}}$$

we have $x_1'^{j_1}x_2'^{j_2} \equiv 0 \pmod{\mathfrak{A}}$ whenever $j_1 + j_2 \geqq M + 1$ or

$$(3.14) \qquad q(x') \equiv 0 \pmod{\mathfrak{A}}$$

where q is any homogeneous polynomial of degree $> M$. Substituting $x' = T^{-1}x^s$ in (3.14) we obtain the desired conclusion.

To prove the first part of Theorem (3.2) it is then enough to show that $P_2 \nmid P_m$ for some $m \leqq N$. For it would then follow that $M(x) \in \mathfrak{P}$ where M is an arbitrary monomial of degree $> N$. This in turn implies that all partial derivatives of $f(x)$ of order $> N$ are zero so that $f(x)$ is a polynomial of degree $\leqq N$. But $P_2 \nmid P_m$ for some $m \leqq N$ is equivalent to saying that the only zero common to

$$(3.15) \qquad P_m(x) = \sum_{j=1}^{N} \mu_j (x \cdot y_j)^m \qquad (1 \leqq m \leqq N)$$

is $x = 0$. Let $b_1, \cdots, b_K$ be the distinct non-zero values of $x \cdot y_j$ $(1 \leqq j \leqq N)$, if there be any such values.

(3.15) may be rewritten as

$$\sum_{j=1}^{K} b_j{}^m \alpha_j \qquad (1 \leqq m \leqq N) \tag{3.16}$$

where $\alpha_j > 0$ $(1 \leqq j \leqq K)$. But the first K equations of (3.16) have a non-zero determinant so that $\alpha_1 = \alpha_2 = \cdots = \alpha_K = 0$, which is a contradiction. Hence

$$x \cdot y_j = 0 \qquad (1 \leqq j \leqq N). \tag{3.17}$$

Since the Y_j's are not all contained in a line through the origin, the rank of (3.17) will be two so that $x = 0$.

If $P_2 \mid P_m$ for some $m < N$, then all solutions of (3.6) are polynomials of degree $< N$. If $P_2 \nmid P_N$ and $P_2 \mid P_m$ for $2 < m < N$, then $\mathfrak{P}_m$ (see Section 2 for definition of p_N and V_N) consists of all $P = P_2 R + c P_N$ where R is homogeneous of degree $N-2$ and c is an arbitrary constant. Thus $\dim \mathfrak{P}_N = (N-1) + 1 = N$ and by Theorem (2.2) $\dim V_N = (N+1) - N = 1$ which means that (3.6) has a polynomial solution of degree N. Hence (3.6) has a polynomial solution of degree N if and only if $P_2 \mid P_m$ $(2 \leqq m \leqq N-1)$; i. e., if and only if

$$P_m(x) = 0 \qquad (1 \leqq m \leqq N-1) \tag{3.18}$$

has a solution $x \neq 0$.

By the Lemma (3.18) is equivalent to

$$x \cdot y_j = \zeta_j \qquad (1 \leqq j \leqq N) \tag{3.19}$$

where $\zeta_1, \cdots, \zeta_N$ are the N N-th roots of a fixed number ζ. Letting $x = u + iv$, this means that

$$\begin{cases} u \cdot y_j = \operatorname{Re}(\zeta_j) \\ v \cdot y_j = \operatorname{Im}(\zeta_j) \end{cases} \qquad (1 \leqq j \leqq N). \tag{3.20}$$

Let A be the matrix whose first row is u and whose second row is v. Since $P_2(x)$ is positive definite, x cannot be a constant multiple of a real vector. This means that u and v are linearly independent vectors so that A is non-singular. (3.6) will thus have a polynomial solution of degree N if and only if (3.7) holds.

We remark that we have not been able to find a sharp bound on the degree of the polynomial solutions of (3.6) when the dimension is greater than two.

4. The case of measures which remain invariant under a finite irreducible o. g. g. r. Let G be a group of orthogonal transformations and suppose that the measure μ is invariant under G. Let $T \in G$. Then

$$P_m(Tx) = \int_Y (Tx \cdot y)^m d\mu(y) = \int_Y (x \cdot T^{-1}y)^m d\mu(y) \tag{4.1}$$

$$= \int_Y (x\cdot y)^m d\mu(Ty) = \int_Y (x\cdot y)^m d\mu(y) = P_m(x) \qquad (1 \leqq m < \infty)$$

Hence $P_m(x)$ is an invariant of the group G for $1 \leqq m < \infty$. We specialize G to be a finite irreducible o. g. g. r. Coxeter [6] has enumerated these groups and has furthermore made the following observation concerning their invariants. Suppose that the group G acts on n variables $x_1, \cdots, x_n$. Then there exist a set of n algebraically independent homogeneous invariants

$$Q_1(x) = \sum_{i=1}^{n} x_i^2, Q_2(x), \cdots, Q_n(x)$$

of respective degrees $d_1, d_2, \cdots, d_n$ $(2 = d_1 < d_2 < \cdots < d_n)$. Every invariant polynomial is a polynomial in $Q_1(x), \cdots, Q_n(x)$ ([14], p. 288). The degrees satisfy the following two relations

$$(4.2) \qquad \begin{cases} \prod_{i=1}^{n} d_i = g, & g = \text{order of } G \\ \sum_{i=1}^{n} d_i - n = r, & r = \text{number of reflections in } G. \end{cases}$$

We quote the following result of Steinberg ([15]) which we will use later on.

THEOREM 4.1. *Let G be a finite o.g.g.r and suppose $R_1, \cdots, R_r$ are the reflections in G with the respective reflecting hyperplanes $L_i(x) = 0$ $(1 \leqq i \leqq r)$. Let $Q_1(x), \cdots, Q_n(x)$ be a basic set of invariants for G. Then the solution space of*

$$(4.3) \qquad Q_i(\partial/\partial x)f = 0 \qquad (1 \leqq i \leqq n)$$

is given by the linear combinations of the partial derivatives of $\prod_{i=1}^{n} L_i(x)$. The dimensionality of the solution space equals the order of G.

Before stating our next result, we require the following.

LEMMA. *Let $M_k = \{x \mid Q_i(x) = 0,\ 1 \leqq i \leqq k\}$. Then M_{k+1} is a proper subset of M_k $(1 \leqq k \leqq n-1)$.*

Proof. The above will certainly be proven if we can show that $\dim M_k > \dim M_{k+1}$. In general we have $\dim M_{k+1} = \dim M_k$ or $\dim M_{k+1} = \dim M_k - 1$ (12, p. 112). Clearly $\dim M_1 = n-1$. Since 0 is the only common complex zero of $Q_1(x), \cdots, Q_n(x)$ ([14], p. 288), we have $\dim M_n = 0$.

Since $\dim M_1 = n-1$, $\dim M_n = 0$, we must indeed have $\dim M_{k+1} = \dim M_k - 1$ $(1 \leqq k \leqq n-1)$. Thus M_{k+1} is properly contained in M_k

and we may choose $\zeta_k = (\zeta_{k1}, \cdots, \zeta_{kn}) \in M_{k-1} - M_k$ $(2 \leqq k \leqq n)$. Consider now the polynomials $P_{d_1}(x), \cdots, P_{d_n}(x)$. Each of these is an invariant and hence a polynomial in $Q_1, \cdots, Q_n$. Since $\deg P_{d_j} = \deg Q_j$ $(1 \leqq j \leqq n)$, Q_j must appear as a linear term in Pd_j. I. e.,

$$P_{d_j}(x) = F_j(Q_1(x), \cdots, Q_{j-1}(x)) + c_jQ_j(x) \qquad (2 \leqq j \leqq n) \tag{4.4}$$

where F_j is a polynomial and c_j a constant.

Substituting ζ_j for x in (4.4) we obtain

$$P_{d_j}(\zeta_j) = \int_Y (\zeta_j \cdot y)^{d_j} d\mu(y) = c_jQ_j(\zeta_j). \tag{4.5}$$

We now stipulate that μ satisfies the following moment condition.

$$\int_Y (\zeta_j \cdot y)^{d_j} d\mu(y) \neq 0 \qquad (2 \leqq j \leqq n) \tag{4.6}$$

Condition (4.6) implies that $c_j \neq 0$ $(2 \leqq j \leqq n)$. Using (4.4) and a straightforward induction on j, we conclude that

$$(P_{d_1}, \cdots, P_{d_n}) = (Q_1, \cdots, Q_n).$$

Since $P_m(x)$ $(1 \leqq m \leqq \infty)$ is an invariant, we have $P_m \in (Q_1, \cdots, Q_n) = (P_{d_1}, \cdots, P_{d_n})$, so that $\mathfrak{P} = (P_{d_1}, \cdots, P_{d_n})$. System (1.2) is thus equivalent to (4.3) if we stipulate condition (4.6). We summarize the abovt discussion as follows.

THEOREM 4.2. *Let G be a finite irreducible o. g. g. r. and let $R_1, \cdots, R_r$ be the reflections of G with respective reflecting hyperplane $L_i(x) = 0$ $(1 \leqq i \leqq r)$. Let μ invariant under G and let it satisfy the moment condition* (4.6). *Then the solution space of* (1.1) *consists of the linear combinations of the partial derivatives of $\prod_{i=1}^{r} L_i(x)$. The dimensionality of the solution space equals the order of G.*

We remark that the last statement of Theorem (4.2) also follows from the corollary to Theorem (2.2).

5. The regular polyhedra. As we had already mentioned in the introduction, the moment condition (4.6) seems rather difficult to check, even for simple measures. In the ensuing discussion we limit ourselves to the case of regular polyhedra. I. e., we wish to determine the class of functions $f(x)$ continuous in a region R such that the value of $f(x)$ at the center of a regular polyhedron similar to a given one equals the average of $f(x)$ with respect to a measure distributed over the polyhedron. As is well known, there are

only three regular polyhedra in n dimensions. These are: the n-dimensional tetrahedron, the n-dimensional octahedron, and the n-dimensional cube ([5], p. 136). With each of these figures we associate the group of orthogonal transformations which leave the figure invariant. The cube and octahedron are duals of each other and are thus associated with the same group. We call the groups which leave the tetrahedron (octahedron) invariant the tetrahedral (octahedral) group. Clearly these two groups are generated by reflections so that the problems of this section are specializations of those considered in Section 4.

For each of the above mentioned figures, there are three natural m. v. p.'s which present themselves.

a) The vertex problem: $\mu(P_i) = 1/N$ $(1 \leqq i \leqq N)$ where $p_1, p_2, \cdots, p_N$ denote the vertices of the polyhedron, $\mu = 0$ elsewhere.

b) The volume problem: μ is the normalized Euclidean measure of the solid polyhedron.

c) The surface problem: μ is the normalized Euclidean measure of the surface of the polyhedron.

It is easy to see that b) and c) are equivalent. We therefore just consider problems a) and b). Walsh [17] has treated the vertex problem in two dimensions. Subsequently, Beckenbach and Reade [1], [2] have treated the volume problem in two dimensions and the vertex problem in three dimensions.

We adopt the following notation. For a given sequence of indices (each $\geqq 0$),

$a = (a_1, \cdots, a_n)$, $|a| = a_1 + \cdots + a_n$, $a! = a_1! \cdots a_N!$, $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

We first state the following lemma used in the sequel.

LEMMA. *Let* $F_k(x) = \sum_{|a|=k} x^a$ *and* $\sigma_k(x) = \sum_{1 \leqq j_1 < j_2 < \cdots < j_k \leqq N} x_{j_1} \cdots x_{j_k}$. *Then* $(F_1, F_2, \cdots, F_n) = (\sigma_1, \sigma_2, \cdots, \sigma_n)$, $(1 \leqq i \leqq n)$.

Proof. We claim that

$$F_i - \sigma_1 F_{i-1} + \sigma_2 F_{i-2} - \cdots + (-1)^i \sigma_i = 0 \qquad (1 \leqq i \leqq n) \tag{5.1}$$

For consider any monomial $x_{i_1}^{j_1} \cdots x_{i_h}^{j_h}$, $j_r > 0$ $(1 \leqq r \leqq h)$, $\sum_{r=1}^{h} j_r = i$, $i_1 < i_2 < \cdots < i_h$. It will occur in $G_i = F_i - \sigma_i F_{i-1} + \cdots + (-1)^i \sigma_i$ with the coefficient

$$1 - \binom{h}{1} + \binom{h}{2} - \cdots + (-1)^h \binom{h}{h} = 0.$$

Hence $G_i = 0$.

The lemma may now be established by induction on i. $(F_1)=(\sigma_1)$ as $F_1=\sigma_1$. Suppose $(F_1,\cdots,F_{i-1})=(\sigma_1,\cdots,\sigma_{i-1})$. We conclude from (5.1) that $F_i\in(\sigma_1,\cdots,\sigma_i)$, $\sigma_i\in(F_1,\cdots,F_i)$ so that $(F_1,\cdots,F_i)=(\sigma_1,\cdots,\sigma_i)$.

We proceed to discuss the m. v. p. for the above mentioned figures.

The n-dimensional tetrahedron. This figure can best be described by imbedding it in $n+1$ dimensions. We assume that the $n+1$ vertices $p_1,\cdots,p_{n+1}$ are given by $p_i=(p_{i1},\cdots,p_{i,n+1})$ where $p_{ii}=n/n+1$, $p_{ij}=-1/n+1$ $(i\neq j)$. Without loss of generality we assume $f\in C^1$; otherwise we first reason with $f_n(x)=\int f(x-v)\delta_n(v)\,dv$ where $\delta_n(v)$ is defined as in Section 3 and then pass to the limit by a standard argument. We furthermore assume that (5.2) $f(x)$ is constant along lines parallel to $e=(1,1,\cdots,1)$. I.e., $Q(\partial/\partial x)\cdot f=0$ where $Q(x)=x_1+x_2+\cdots+x_{n+1}$. The m. v. p. (1.1) is then equivalent to

$$P_m(\partial/\partial x)\cdot f=0 \qquad (2\leqq m<\infty) \tag{5.3}$$

where $P_m(x)=\int_Y(x\cdot y)^m d\mu(y)$. Let $z=y+e/n+1$. Then

$$P_m(x)=\int_Z(x\cdot z-x\cdot e/n+1)^m d\eta(z)\equiv\int_Z(x\cdot z)^m d\eta(z)\ (\mathrm{mod}\,Q(x))$$

where η is the normalized Euclidean measure of the set $Z: z_1+\cdots+z_{n+1}=1$, $z_i\geqq 0$ $(1\leqq i\leqq n+1)$. Thus

$$P_m(x)\equiv 1/n+1[x_1{}^m+\cdots+x_{n+1}{}^m] \qquad (\mathrm{mod.}\,Q) \tag{5.4}$$

for the vertex problem and

$$P_m(x)\equiv\sum_{|a|=m}\frac{m!\int_Z z^a d\eta(z)}{a!}\,x^a \qquad (\mathrm{mod}\,Q) \tag{5.5}$$

for the volume problem. Now

$$\int_Z z^a d\sigma(z)=\frac{a!\sqrt{n+1}}{(m+n)!}, \tag{5.6}$$

where σ is the Euclidean measure of Z. ([19], p. 252) Hence

$$P^m(x)\equiv\frac{m!n!}{(m+n)!}\,F_m(x). \tag{5.7}$$

We conclude from the lemma that $(Q,P_2,\cdots,P_{n+1})=(\sigma_1,\sigma_2,\cdots,\sigma_{n+1})$. Since $P_m(x)$ $(z\leqq m<\infty)$ is symmetric in $x_1,x_2,\cdots,x_{n+1}$ we have $P_m\in(\sigma_1,\cdots,\sigma_{n+1})=(Q,P_2,\cdots,P_{n+1})$. Letting $\mathfrak{P}$ denote the ideal generated by Q and P_m $(z\leqq m<\infty)$, we have $\mathfrak{P}=(\sigma_1,\cdots,\sigma_{n+1})$. As $\sigma_1,\sigma_2,\cdots,\sigma_{n+1}$ form a basic set of invariants for the tetrahedral group,

we conclude from Theorem (4.1) that the solution space of (1.1), both for the vertex and volume problems, is given by the linear combinations of the partial derivatives of $\prod_{1\leqq i<j\leqq n+1}(x_i-x_j)$.

The n-dimensional octahedron. We assume that this figure has the vertices $(1,0,\cdots,0),(0,1,0,\cdots,0),\cdots,(0,\cdots,0,1)$. The vertex problem is treated in [9] and we conclude it here for the sake of completeness. We have

$$\begin{cases} P_{2m}(x)=1/2n[x_1^{2m}+\cdots+x_n^{2m}] & (1\leqq m<\infty) \\ P_{2m-1}(x)=0 \end{cases} \tag{5.8}$$

for the vertex problem while

$$\begin{cases} P_{2m}(x)=\dfrac{2^n}{V_n}\displaystyle\sum_{|a|=m}\dfrac{(2m)!\displaystyle\int_Y y^{2a}dy}{(2a)!}\,x^{2a} \\ P_{2m-1}(x)=0 \qquad (1\leqq m<\infty) \end{cases} \tag{5.9}$$

for the volume problem. Here Y denotes the set $\sum_{i=1}^{N}y_i\leqq 1,\ y_i\geqq 0\ (1\leqq i\leqq n)$ and V_n denotes the volume of the octahedron. Now

$$\int_Y y^{2a}dy=\frac{(2a)!}{(2m+n)!}. \tag{5.10}$$

([19], p. 252) Hence

$$P_{2m}(x)=\frac{(2m)!n!}{(2m+n)!}\,F_m(x) \qquad (1\leqq m<\infty) \tag{5.11}$$

where $x^2=(x_1^2,\cdots,x_n^2)$.

It follows from the lemma that $(P_2,P_4,\cdots,P_{2n})=(\sigma_1(x^2),\cdots,\sigma_n(x^2))$. Since $P_m(x)$ $(1\leqq m<\infty)$ is a symmetric function of $x_1^2,\cdots,x_n^2$, $\mathfrak{P}=(\sigma_1(x^2),\cdots,\sigma_n(x^2))$, $\sigma_1(x^2),\sigma_2(x^2),\cdots,\sigma_n(x^2)$ form a basic set of invariants for the octahedral group and it follows from Theorem (4.1) that the solutions of (1.1) are the linear combinations of the partial derivatives of $x_1,x_2\cdots x_n\prod_{1\leqq i<j<n}(x_i^2-x_j^2)$.

The n-dimensional cube. We pursue the method outlined in Section 4. We choose $\sigma_1(x^2),\cdots,\sigma_n(x^2)$ as our basic set of invariants. We furthermore choose

$$\zeta_k=(e^{2\pi i/k},e^{2\pi i/k},\cdots,e^{\pi i},0,\cdots,0) \qquad (2\leqq k\leqq n). \tag{5.12}$$

Let $\xi_j=e^{2\pi ij/k}$, $(1\leqq j\leqq k)$. Condition (4.8) becomes

$$\alpha_k=\sum_{|a|=k}\frac{(2k)!}{(2a)!}\,\xi^a\neq 0 \tag{5.13}$$

for the vertex problem and

$$\beta_k = \sum_{|a|=k} \frac{(2k)!}{(2a+1)!} \xi^a \neq 0 \tag{5.14}$$

for the volume problem.

We have not been able to prove either (5.13) or (5.14) and we limit ourselves to the following partial result concerning α_k and β_k.

THEOREM 5.1. *Let $2k-1=p^s$ where p is a prime. Then* (5.13) *holds. Let $2k+1=p$ where p is a prime. Then* (5.14) *holds.*

Proof. We have

$$\alpha_k = k + \sum{}' \frac{(2k)!}{(2a)!} \xi^a \tag{5.15}$$

where $\sum'$ includes only those a's for which $a_j < k$ $(1 \leqq j \leqq k)$ α_k is an algebraic integer and since it is symmetric in the ξ's, it is a rational integer. we will show that $p \mid \sum'$. Since $p \nmid k$, $p \nmid \alpha_k$ and thus $\alpha_k \neq 0$. To show that $p \mid \sum'$, we consider the highest power M of p which divides $\dfrac{(2k)!}{(2a)!}$ where $a_i < k$ $(1 \leqq i \leqq k)$. This number M is given by ([13], p. 90)

$$M = \sum_{j=0}^{s} \{[2k/p^j] - ([2a_1/p^j] + \cdots + [2a_k/p^j])\} \tag{5.16}$$

Now

$$[2k/p^s] - ([2a_1/p^s] + \cdots + [2a_k/p^s]) = 1$$

while

$$[2k/p^s] - ([2a_1/p^s] + \cdots + [2k_k/p^s]) \geqq 0 \qquad (1 \leqq j < s).$$

Hence $M \geqq 1$, which means that p divides every term in $\sum'$ so that $p \mid \sum'$.

We prove the second part of Theorem (5.1) in an analogous fashion. We have

$$(2k+1)(2k+2)\cdots 3k\beta_k = (2k+2)\cdots 3k\cdot k + \sum{}' \frac{(3k)!}{(2a+1)!}\xi^a$$

where $\sum'$ includes only those a's for which $a_j < k$ $(1 \leqq j \leqq k)$. The same reasoning as before shows that $\sum'$ is an integer divisible by p so that $\beta_k \neq 0$.

If $\alpha_k \neq 0$ $(2 \leqq k \leqq n)$, then it follows from Theorem (4.2) that the solution space of the vertex problem is given by the linear combinations of the partial derivatives of $x_1 \cdots x_n \prod^{N} (x_i^2 - x_j^2)$. The same result holds for the volume problem provided $\beta_k \neq 0$ $(2 \leqq k \leqq n)$. In view of Theorem (5.1) $\alpha_k \neq 0$ for $k \leqq 7$, $\beta_k \neq 0$ for $k \leqq 3$. Hence, the solution space of the vertex problem is known explicitly for $n \leqq 7$ while the one of the volume problem is known explicitly for $n \leqq 3$.

These bounds can be improved by using slightly different congruence arguments but we have not been able to answer the general question. In any case, it follows from Sections 3 and 4 that the solution space, both for the vertex and volume problems, is finite dimensional.

We now discuss the space of functions with the rotated mean value property (r. m. v. p.)

$$(5.17)\qquad f(x)=\int_Y f(x+ty)\,d\mu(Ty)\ \text{ for } x\in R,\ 0<t<\epsilon_x$$

where T stands for an arbitrary rotation and μ stands for any of the six measures described above. The same method as in [8] shows that (5.17) is equivalent to the infinite system of equations

$$(5.18)\qquad P_m(T\partial/\partial x)\cdot f=0 \qquad (1\leqq m<\infty)$$

where $P_m(x)=\int_Y (x\cdot y)^m d\mu(y)$, T being an arbitrary rotation. In particular $P_2(x)=\left(\int_Y y_1^2 d\mu(y)\right)\cdot\sum_{i=1}^N x_i^2$.

It follows from the work of Brelot and Choquet (3) that if $f(x)$ satisfies

$$(5.19)\qquad \begin{cases}\sum_{i=1}^N \partial^2 f/\partial x_i^2=0\\ S(T\partial/\partial x)\cdot f=0\end{cases}$$

where S is a homogeneous polynomial of degree $m>2$, T is an arbitrary rotation, and $\sum_{i=1}^N x_i^2 \nmid S$, then $f(x)$ is a harmonic polynomial of degree $<m$. Conversely, and such harmonic polynomial satisfies (5.19). Suppose then that m_0 is the smallest integer >2 such that $P_2\nmid P_{m_0}$. The solution space of (5.17) will consist of harmonic polynomials of degree $<m_0$. It follows that if μ is the vertex or volume measure of the cube or octahedron, then the solutions of (5.17) are the harmonic polynomials of degree $\leqq 3$.

In case of the tetrahedron, we introduce an orthogonal transformation $u=Tx$ such that $u_{n+1}=1/\sqrt{n+1}\,(x_1+\cdots+x_{n+1})$. Then $\partial/\partial u=T\partial/\partial x$ and (5.19) is in this case equivalent to

$$(5.10)\qquad \begin{cases}\partial f/\partial u_{n+1}=0\\ \sum_{j=1}^{n+1}\partial^2 f/\partial u_j^2=0\\ S_m(T\partial/\partial u)\cdot f=0 \qquad (3\leqq m\leqq n+1)\end{cases}$$

where T is an arbitrary rotation in the variables $u_1,\cdots,u_n$ and

$$S_m(u) = Q_m(T^{-1}u), \qquad Q_m(x) = \sum_{i=1}^{n+1} x_i^m.$$

Now $Q_3(x) \notin (\sum_{i=1}^{n+1} x_i, \sum_{i=1}^{n+1} x_i^2)$ since $Q_3(\xi_3) \neq 0$ while $Q_1(\xi_3) = Q_2(\xi_3) = 0$; here $\xi_3 = (1, e^{2\pi i/3}, e^{4\pi i/3}, 0, \cdots, 0)$. Hence $S_3(u) \notin (u_{n+1}, \sum_{j=1}^{n+1} u_j^2) = (u_{n+1}, \sum_{j=1}^{n} u_j^2)$. Furthermore $S_3(u) \equiv S_3'(u_1, \cdots, u_n) \bmod (u_{n+1}, \sum_{i=1}^{n} u_i^2)$ where S_3' is not divisible by $\sum_{i=1}^{n} u_i^2$. Since $S_3'(T\partial/\partial u) \cdot f = 0$ for all n-dimensional rotations in $u_1, u_2, \cdots, u_n$, we conclude that f is a harmonic polynomial of degree 2. Conversely any such f satisfies (5.17).

We tabulate the results of this section as follows:

Figure	Solution Space to m. v. p. Problem	Solution Space to r. m. v. p. Problem
Tetrahedron (vertex)	Linear combinations of partial derivatives of $\prod_{1 \leq i < j \leq n+1} (x_i - x_j)$	Harmonic polynomials of degree ≤ 2
" (volume)	"	"
Octahedron (vertex)	Linear combinations of partial derivatives of $x_1 \cdots x_n \prod_{1 \leq i < j \leq n} x_i^2 - x_j^2$	Harmonic polynomials of degree ≤ 3
(volume)	"	"
Cube (vertex)	?	"
(volume)	?	"

6. The dodecahedron and the icosahedron. We assume that the icosahedron has the following orientation. One of the vertices is at $N = (0, 0, 1)$, the x_3-axis being an axis of symmetry. The x_1x_3 plane bisects one of the faces which has N for a vertex. The vertices of the dodecahedron are the centroids of the faces of the icosahedron. Let G be the icosahedral group; i. e., the group of orthogonal transformations which leave the icosahedron (dodecahedron) invariant. Goursat (11) has computed the following set of basic invariants for G.

$$(6.1) \quad \begin{cases} Q_2(x) = x_1^2 + x_2^2 + x_3^2 \\ Q_6(x) = x_3^6 - 5(ss_0)x_3^4 + 5(ss_0)^2 x_3^2 - (s^5 + s_0^5)x_3 \\ Q_{10}(x) = x_3^{10} - 35(ss_0)x_3^8 + 260(ss_0)^2 x_3^6 - 175(ss_0)^3 x_3^4 \\ \qquad + 25(ss_0)^4 x_3^2 + (s^5 + s_0^5)[123 x_3^5 - 90(ss_0)x_3^2 \\ \qquad + 15(ss_0)^2 x_3] + (s^5 + s_0^5)^2. \end{cases}$$

Here, $s=x_1+ix_2$, $s_0=x_1-ix_2$. We choose $\zeta_3=(1,i,0)$. We will show that $\mathfrak{P}=(Q_2,Q_6,Q_{10})$ for the vertex problems associated with the icosahedron and the dodecahedron. We treat in detail the case of the icosahedron, the argument for the dodecahedron being almost identical. As pointed out in Section 4, it suffices to show that $P_6\notin(P_2)$, $P_{10}\notin(P_2,P_6)$. Let T be a similarity transformation which carries the vertices of the above mentioned icosahedron into the one with the vertices $(\pm m,\pm 1,0)$, $(0,\pm m,\pm 1)$, $(\pm 1,0,\pm m)$ where $m=\dfrac{1+\sqrt{5}}{2}$ ([5], p. 52). Denoting the vertices of the original icosahedron by y_j $(1\leqq j\leqq 12)$ we have

$$
\begin{aligned}
12P_6(T^{-1}x)&=\sum_{j=1}^{12}(T^{-1}x,y_j)^6=\frac{1}{\det^2 T}\sum_{j=1}^{12}(x,Ty_j)^6\\
(6.2)\qquad &=\frac{1}{\det^2 T}\Big[(1+m^6)\sum_{k=1}^{3}x_k^6+15m^4\sum_{k=1}^{3}x_k^4x_{k+1}^2+15m^2\sum_{k=1}^{3}x_k^2x_{k+1}^4\Big]
\end{aligned}
$$

(Here $x_4=x_1$.) Eliminating x_3 between $x_1^2+x_2^2+x_3^2$ and $(12\det^2 T)P_6(T^{-1}x)$ we find that

$$
(6.3)\qquad (12\det^2 T)(P_6(T^{-1}x)=R(x_1,x_2)\quad(\operatorname{mod} x_1^2+x_2^2+x_3^2)
$$

where the coefficients of x_1^6 in $R=15(m^4-m^2)\neq 0$ so that $R\neq 0$. Thus $P_2(x)\nmid P_6(T^{-1}x)$ or, equivalently, $P_2(x)\nmid P_6(x)$. Now

$$
(6.4)\qquad 12P_{10}(\zeta_3)=\sum_{j=1}^{12}\eta_j^{10}
$$

where $\eta_j=y_{j1}+y_{j2}i$, $(1\leqq j\leqq 10)$.

The η_j's are the projections of the vertices upon the x_1x_2 plane. Two of these will be 0, while the others are the ten tenth roots of a positive number r. Thus

$$
(6.5)\qquad 12P_{10}(\zeta_3)=10r\neq 0.
$$

It follows from the discussion in Section 4 that $P_{10}\notin(P_2,P_6)$. The argument for the dodecahedron proceeds along similar lines. We just mention that T is to be replaced by the similarity transformation T' which takes the vertices y_j $(1\leqq j\leqq 20)$ into the vertices $(\pm[m+1],0,\pm 1)$ $(\pm 1,\pm[m+1],0)$, $(0,\pm 1,\pm[m+1])$, $(\pm m,\pm m,\pm m)$, ([5], p. 52).

We therefore conclude from Theorem (4.1) that for both of the above mentioned figures, the solution space consists of the linear combinations of partial derivatives of $\prod_{j=1}^{15}L_j(x)$ where $L_j(x)=0$ $(1\leqq j\leqq 15)$ denote the fifteen reflecting planes of the icosahedral group.

YESHIVA UNIVERSITY.

REFERENCES.

[1] E. F. Beckenbach and M. Reade, "Mean values and harmonic polynomials," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 230-238.

[2] E. F. Beckenbach and M. Reade, "Regular solids and harmonic polynomials," *Duke Mathematical Journal*, vol. 12 (1945), pp. 629-644.

[3] M. Brelot and G. Choquet, "Polynomes harmoniques et polyharmoniques," *Second Colloque Sur Les Equations Aux Dérivées Partielles*, Bruxelles, March 1954, pp. 44-66.

[4] W. Brödel, "Funktionen mit Gaussicher Mittelwerteigenschaft für Konvexen Kurven und Bereiche," *Deutsche Mathematik*, vol. 4 (1939), pp. 3-15.

[5] H. S. M. Coxeter, *Regular Polytopes*, Methuen & Co., Ltd., London, 1948.

[6] ———, "The product of the generators of a finite group generated by reflections," *Duke Mathematical Journal*, vol. 18 (1951), pp. 765-782.

[7] E. Fischer, "Uber Algebraische Modulsysteme und Lineare Homogene Partiel Differentialgleichungen mit Konstaten Koeffizienten," *Journal für die Reine und Angewandte Mathematik*, vol. 140 (1911), pp. 48-81.

[8] L. Flatto, "Functions with a mean value property," *Journal of Mathematics and Mechanics*, vol. 10 (1961), pp. 11-18.

[9] A. Friedman and W. Littman, "Functions satisfying the mean value property," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 167-180.

[10] A. M. Garsia, "A note on the mean value property," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 181-186.

[11] E. Goursat, "Études des Surfaces qui admettent les Plans de Symmetrie d'un Polyèdre Régulier," *Annales Scientifique de l'École Normale Supérieure* (3), vol. 4 (1887), pp. 158-200.

[12] W. Gröbner, *Moderne Algebraische Geometrie*, Springer-Verlag, Wien und Innsbruck, 1949.

[13] W. J. LeVeque, *Topics in Number Theory*, vol. I, Addison-Wesley, Reading, Mass.

[14] G. C. Shephard and J. A. Todd, "Finite unitary reflection groups," *Canadian Journal of Mathematics*, vol. 6 (1954), pp. 274-304.

[15] R. Steinberg, "Differential equations invariant under finite reflection groups," *Notices of the American Mathematical Society*, vol. 9 (February, 1962).

[16] B. L. van der Waerden, *Modern Algebra*, vol. I, Frederick Ungar Publishing Company, New York ,1950.

[17] J. L. Walsh "A mean value theorem for polynomials and harmonic polynomials," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 923-930.

[18] H. Weyl, *The Classical Groups*, Princeton University Press, Princeton, N. J., 1946.

[19] Whittaker and Watson, *A Course of Modern Analysis*, Cambridge University Press, 1915.

ON THE INDEX THEOREM.* [1]

By Leif-Norman Patterson.

1. Introduction. This paper purports to present an account of the general index problem in Riemannian geometry from a geometric point of view, to point out some of the most important index theorems, and to study some continuity questions related to the index problem. The results obtained are: 1. All self-adjoint boundary-value problems arising in Riemannian geometry can be interpreted geometrically in terms of separate endmanifolds. 2. The subsequent extension of an index theorem [2] of Ambrose [2] to all such problems, in particular to that associated with periodic geodesics. 3. The continuous dependence (with some exceptions) of the Ambrose conjugate points on the boundary conditions and on the Riemannian metric on the manifold.

We thank Professor W. Ambrose for many helpful conversations. Our notation will follow closely that in two papers of his [1], [2], and it will be assumed the reader has some familiarity with these. Attention is called to the fact that Bott [4] and Edwards [7] have developed a powerful approach to the index problem through topological methods.

We begin by recalling some standard facts regarding product manifolds, which are basic for the following theory. Let M_1 and M_2 be Riemannian manifolds with metric tensor-fields $\langle\,,\rangle_1$ and $\langle\,,\rangle_2$. $N = M_1 \times M_2$ becomes a Riemannian manifold when we define $\langle x, y\rangle = \langle d\pi_1 x, d\pi_1 y\rangle_1 + \langle d\pi_2 x, d\pi_2 y\rangle_2$, where π_1 and π_2 are the natural projections on the factors. The map $x \to (d\pi_1 x, d\pi_2 x) : TN \to TM_1 \times TM_2$ is a diffeomorphism and these two spaces are indeed often identified. If $X = (X_1, X_2)$ is a differentiable curve in $TM_1 \times TM_2$, that is X is a vectorfield along a curve σ in N, and if X' is the covariant derivative of X along σ, then $X' = (X_1', X_2')$ where X_i' is the covariant derivative in M_i of X_i along the projection $\sigma_i = \pi_i \circ \sigma$ of σ. σ is a geodesic in N if and only if both its projections σ_1 and σ_2 are geodesics in M_1 and M_2 respectively. $X = (X_1, X_2)$ is a Jacobifield along the geodesic σ in N if and only if X_1 and X_2 are Jacobifields along σ_1 and σ_2. Finally,

* Received December 20, 1962.

[1] Part of this work was supported by NSF grant G21938.

[2] T. Ōtsuki has recently found and corrected errors in the original proof by Ambrose.

suppose the geodesic σ passes through $m=(m_1,m_2)$ and $n=(n_1,n_2)$, and let $\nu_\sigma(m,n)$ be the order of m as a conjugate point of n along σ. Then $\nu_\sigma(m,n)=\nu_{\sigma_1}(m_1,n_1)+\nu_{\sigma_2}(m_2,n_2)$. Note that if $M_1=M_2$, then the diagonal Δ of $N=M_1\times M_1$ is a totally geodesic submanifold of N, and hence its second fundamental form is 0.

2. The index form and geometric boundary conditions. Let M be a complete Riemannian manifold, and B a submanifold of $M\times M$. $\Omega(B)$ is the space of broken C^∞ curves in M parametrised by reduced arclength over the interval $[0,1]$ and subject to the constraint B; i.e. $\sigma\in\Omega(B)$ implies $(\sigma(0),\sigma(1))\in B$. On $\Omega(B)$ the usual length function L is defined by $L(\sigma)=\int_0^1|\sigma_*|=|\sigma_*|$. $\sigma\in\Omega(B)$ is a critical point of L if and only if σ is a geodesic in M such that its "lift" g in $M\times M$ hits B perpendicularly at $g(0)$. (Proposition 2.1). g is the curve in $M\times M$ defined by $g(t)=(\sigma(t),\sigma(1-t))$. That is, $g=(\sigma,\sigma^-)$. At a critical point σ of L Synge's formula gives information about the second variation of L in terms of a bilinear symmetric form I_B which is a sum of two forms: an integral which does not depend on B, and an endpoint form which depends on $S^B{}_{g_*(0)}$. (Proposition 2.2). $S^B{}_{g_*(0)}$ is the second fundamental form of B in $M\times M$ relative to the direction $g_*(0)$, and I_B is defined on $\Xi\times\Xi$, where Ξ is the linear space of all broken C^∞ vectorfields X along σ such that $(X(0),X(1))\in B_{g(0)}$. (These are transversal vectorfields generated by curves in $\Omega(B)$ through σ. Intuitively, I_B is the Hessian of L at σ.)

The index problem consists in determining the index of I_B (the dimension of a maximal subspace on which I_B is negative definite) in terms of dimensions of certain spaces of broken Jacobifields; or, equivalently, in terms of a discrete set of points along the geodesic, called conjugate points, and certain other data which depend on the local structure of B at $g(0)$. The relevant local structure of B at $g(0)$ is contained in the tangent space $B_{g(0)}$ and the second fundamental form $S^B{}_{g_*(0)}$, and from these data the notion of a self adjoint boundary condition is abstracted (Definition 4.5 I). A vector field $\tilde{X}$ along g in $M\times M$ *satisfies* the boundary condition B *at* $g(0)$ if and only if $\tilde{X}(0)\in B_{g(0)}$ and $\tilde{X}'(0)-S^B{}_{g_*(0)}\tilde{X}(0)\perp B_{g(0)}$. Here $\tilde{X}'$ denotes the covariant derivative of $\tilde{X}$ along g. The boundary condition thus described is the one-endpoint condition in the manifold $M\times M$ associated with the submanifold B and the geodesic g perpendicular to B at $g(0)$, but it represents a boundary condition in M involving both endpoints of σ in the following sense: If X is a vectorfield along σ in M, then the "lift" $\tilde{X}$ of X is the vector-

field along g in $M \times M$ defined by $\bar{X} = (X, X^-)$, that is $\bar{X}(t) = (X(t), X(1-t))$. X is said to *satisfy* the (self adjoint) boundary condition B if and only if its lift $\bar{X}$ satisfies the boundary condition B at $g(0)$ in the above sense.

If $B = S \times T$ where S and T are submanifolds of M, the associated index problem is referred to as the separated endmanifold case, and X satisfies $B = S \times T$ if and only if X satisfies S at $\sigma(0)$ and T at $\sigma(1)$. If $B = \Delta$ (the diagonal of M) and σ is periodic, then the associated index problem is referred to as the periodic or closed geodesic case.

Next observe that if g is the lift of a curve $\sigma \in \Omega(B)$, then $g(\frac{1}{2}) \in \Delta$, and all the information about σ is contained in the piece of g from 0 to $\frac{1}{2}$. If σ is also a critical point of L, then g hits both of the submanifolds B and Δ of $M \times M$ perpendicularly at $g(0)$ and $g(\frac{1}{2})$ respectively. If $\bar{\Omega}(B, \Delta)$ is the space of curves from B to Δ parametrised on $[0, \frac{1}{2}]$ by reduced arc length, then g (restricted to $[0, \frac{1}{2}]$) is a critical point of the length function $\bar{L}$ on $\bar{\Omega}(B, \Delta)$ and Synge's formula gives the second variation of $\bar{L}$ at g in terms of the bilinear form $I_{B\Delta}$ $(= I_{B \times \Delta})$ on $\bar{\Xi} \times \bar{\Xi}$. Here $\bar{\Xi}$ is the linear space of all broken C^∞ vectorfields along g, starting tangent to $B_{g(0)}$ and ending tangent to $\Delta_{g(\frac{1}{2})}$.

The question to be answered in this section is: What is the relationship between I_B and $I_{B\Delta}$? Geometrically the spaces $\Omega(B)$ and $\bar{\Omega}(B, \Delta)$ are closely related, there being a natural injection of the former into the latter, and it is therefore not surprising that the index of I_B equals that of $I_{B\Delta}$. The importance of this result lies in the fact that all geometric boundary problems (B, σ) in M as described above, can be lifted to equivalent problems $(B \times \Delta, g)$ in $M \times M$ which are of the separate end manifold type.

Even though $B = \{\text{one point}\}$ is not strictly a submanifold of $M \times M$, the theorems apply to this case also, when the tangent bundle TB is interpreted as consisting of the 0-element only.

We now proceed to establish the result quoted above, and for the sake of completeness include

PROPOSITION 2.1. *$\sigma \in \Omega(B)$ is a critical point of L if and only if σ is a geodesic and $g_*(0) \perp B_{g(0)}$, where $g = (\sigma, \sigma^-)$.*

PROPOSITION 2.2. *Let $\sigma \in \Omega(B)$ be a critical point of L. If $X \in \Xi$ $(= \Omega(B)_\sigma)$, if $\alpha : [c, d] \to \Omega(B)$ is a broken C^∞ curve tangent to X at $\sigma = \alpha(c)$, and if $f = L \circ \alpha$, then*

$$(2.1) \qquad f''(c) = (1/k) \int_0^1 \Gamma(X^\perp, X^\perp) + (1/k) \langle S^B{}_{g_*(0)} \bar{X}(0), \bar{X}(0) \rangle.$$

Here $X^\perp$ is the component of X perpendicular to σ_*, $k = |\sigma_*|$, and $\Gamma(X,Y) = \langle X', Y'\rangle - \langle R_{\sigma_* X}\sigma_*, Y\rangle$ where R is the curvature tensor in M.

Proof of Proposition 2.1. We refer to the calculations in [1] regarding broken C^∞ curves in $\Omega(B)$. Precisely, $\alpha: [c,d] \to \Omega(B)$ is called a broken C^∞ curve in $\Omega(B)$ if the map $(t,s) \to \alpha(s)(t): [0,1] \times [c,d] \to M$, also denoted by α, is continuous, and such that for a finite sequence

$$0 < t_1 < t_2 < \cdots < t_r < 1$$

the mappings $\alpha \mid [t_i, t_{i+1}] \times [c,d]$ are C^∞ into M. To demand that α be a curve in $\Omega(B)$ implies furthermore that $(\alpha(0,\cdot), \alpha(1,\cdot))$ is a curve in B. Define a vectorfield along σ in Ξ by $X(t) = \alpha^t{}_*(c)$. Then $dL(X) = f'(c)$ by definition, and the formula below will show that $f'(c)$ depends only on X, not on the whole rectangle α. (α, considered as a curve in $\Omega(B)$, is said to be tangent to X.)

Under the only assumption that α_{c^*}. $(=\sigma_*)$ has constant length, which certainly holds because of the reduced arc length parametrisation, the following formula is derived in [1] via the first structure equation, for any broken C^∞ rectangle α in M. $(dL(X) =)$

$$\begin{aligned} f'(c) = (1/k)\langle\sigma_*(1), X(1)\rangle &- (1/k)\langle\sigma_*(0), X(0)\rangle \\ &+ (1/k)\Sigma(\langle\sigma_*(t_i^-), X(t_i)\rangle - \langle\sigma_*(t_i^+), X(t_i)\rangle) \\ &- (1/k)\int_0^1 \langle\sigma_*', X\rangle. \end{aligned}$$

The two first terms on the right hand side can be combined to yield $-(1/k)\langle g_*(0), \bar{X}(0)\rangle$, when we denote the metric tensor in $M \times M$ also by $\langle\,,\rangle$. (We have used here the fact that

$$g_*(t) = (\sigma_*(t), -\sigma_*(1-t)) = (\sigma_*, -\sigma_*^-)(t).)$$

If $dL = 0$ at σ, then $f'(c) = 0$ for all X (f depends on X). As in [1], by making appropriate choices for X, we conclude 1. σ_* is smooth, 2. $g_*(0) \perp B_{g(0)}$ and 3. $\sigma_*' = 0$, which imply σ is geodesic with lift g perpendicular to B at $g(0)$.

Conversely, the same formula shows that if σ satisfies these conditions, then $dL = 0$ at σ. (Proposition 2.1 represents a transversality condition in the calculus of variations.)

Proof of Proposition 2.2. We assume σ is a critical point of L. α will be a rectangle as in the previous proof. Under the only assumption that σ

is a geodesic in M, the following formula is developed in [1, formula 3.8] via the second structure equation,

$$f''(c) = (1/k)\int_0^1 \Gamma(X^\perp, X^\perp) + (1/k)\langle \alpha_1(1,c), \alpha_2(1,c)\alpha_2\rangle \\ -(1/k)\langle \alpha_1(0,c), \alpha_2(0,c)\alpha_2\rangle.$$

Here $d\alpha(\alpha_1(\cdot,c)) = \sigma_*(\cdot)$ and $d\alpha(\alpha_2(\cdot,c)) = X(\cdot)$. Along the curve $\alpha(0,\cdot)$ we have a vectorfield Z_0, namely the field of tangentvectors to $\alpha(0,\cdot)$. Similarly along $\alpha(1,\cdot)$ we have Z_1. In particular, since $(\alpha(0,\cdot),\alpha(1,\cdot))$ lies in B, $\bar{Z} = (Z_0, Z_1)$ is tangent to B. Since $\langle \alpha_1(0,c), \alpha_2(0,c)\alpha_2\rangle = \langle \sigma_*(0), Z_0(0)Z_0\rangle$, where $Z_0(0)Z_0$ is the covariant derivative of Z_0 in the direction $Z_0(0)$, and since covariant differentiation in a product manifold corresponds to covariant differentiation "in each of the factors," we obtain $f''(c) = (1/k)\int_0^1 \Gamma(X^\perp, X^\perp) - (1/k)\langle \bar{Z}(0)\bar{Z}, g_*(0)\rangle$. But if $g_*(0)$ is a vector perpendicular to B and $\bar{Z}(0)\bar{Z}$ the covariant derivative in $M\times M$ of $\bar{Z}$ in a direction tangential to B, then by a standard fact for second fundamental forms $\langle g_*(0), \bar{Z}(0)\bar{Z}\rangle = \langle -S^B{}_{g_*(0)}\bar{Z}(0), \bar{Z}(0)\rangle$. And observing that $\bar{Z}(0) = (Z_0(0), Z_1(0)) = (X(0), X(1)) = \bar{X}(0)$ we get

$$f''(c) = (1/k)\int_0^1 \Gamma(X^\perp, X^\perp) + (1/k)\langle S^B{}_{g_*(0)}\bar{X}(0), \bar{X}(0)\rangle,$$

which is the desired result.

PROPOSITION 2.3. [7, 2] *Every real self adjoint boundary condition at* $g(0) = (\sigma(0), \sigma(1))$ [*for which* σ *is critical*], *arises from a submanifold* B *of* $M\times M$ [*intersecting* g *perpendicularly at* $g(0)$].

Note. If $\sigma\colon [0,1]\to M$ is a closed, but not smoothly closed (periodic) geodesic, then σ is not a critical point of L in $\Omega(\Delta)$, because $g_*(0)$ is not perpendicular to $\Delta_{g(0)}$.

In the rest of this section we shall adopt the following notation. B will denote a fixed submanifold in $M\times M$, σ a geodesic in M with lift $g = (\sigma, \sigma^-)$ perpendicular to B at $g(0)$. g will be regarded restricted to $[0,\frac{1}{2}]$. Note that $g_*(\frac{1}{2})$ is perpendicular to $\Delta_{g(\frac{1}{2})}$. We consider these linear spaces (over the reals):

$$\Xi\ (= \tilde{\Omega}(B,\Delta)_g) = \{\mathfrak{X} \mid \mathfrak{X} \text{ is a broken } C^\infty \text{ vectorfield along } g \\ (\text{on } [0,\tfrac{1}{2}]) \text{ and } \mathfrak{X}(0)\in B_{g(0)}, \mathfrak{X}(\tfrac{1}{2})\in \Delta_{g(\frac{1}{2})}\}$$

$$\Xi^\perp = \{\mathfrak{X}\in\Xi \mid \mathfrak{X}\perp g_*\}$$

$\bar{\Xi}^0 = \{\mathfrak{X} \in \bar{\Xi} \mid \mathfrak{X} \parallel g_*\}$, $\bar{\Xi} = \bar{\Xi}^\perp + \bar{\Xi}^0$ (direct sum).

$\Xi = \{X \mid X$ is a broken C^∞ vectorfield along σ and $(X(0), X(1)) \in B_{g(0)}\}$

$\Xi^\perp = \{X \in \Xi \mid X \perp \sigma_*\}$.

Corresponding to Synge's formula of Proposition 2.2 we define two index forms I_B and $I_{B\Delta}$ on $\Xi \times \Xi$ and $\bar{\Xi} \times \bar{\Xi}$ respectively by

$$(2.1)\quad I_B(X,Y) = (1/k)\int_0^1 \Gamma(X^\perp, Y^\perp) + (1/k)\langle S^B{}_{g_*(0)}\bar{X}(0), \bar{Y}(0)\rangle$$

$$(2.2)\quad I_{B\Delta}(\mathfrak{X},\mathfrak{Y}) = (k\sqrt{2})^{-1}\int_0^{\frac{1}{2}} \bar{\Gamma}(\mathfrak{X}^\perp, \mathfrak{Y}^\perp) + (k\sqrt{2})^{-1}(\langle S^B{}_{g_*(0)}\mathfrak{X}(0), \mathfrak{Y}(0)\rangle - \langle S^\Delta{}_{g_*(\frac{1}{2})}\mathfrak{X}(\tfrac{1}{2}), \mathfrak{Y}(\tfrac{1}{2})\rangle)$$

$$= (k\sqrt{2})^{-1}\int_0^{\frac{1}{2}} \bar{\Gamma}(\mathfrak{X}^\perp, \mathfrak{Y}^\perp) + (k\sqrt{2})^{-1}(\langle S^B{}_{g_*(0)}\mathfrak{X}(0), \mathfrak{Y}(0)\rangle)$$

where $X, Y \in \Xi$, $k = |\sigma_*|$, $|g_*| = k\sqrt{2}$ and $\mathfrak{X}, \mathfrak{Y} \in \bar{\Xi}$. Γ, $\bar{\Gamma}$ are as in Proposition 2.2.

LEMMA 2.4. a) *The lift* $\bar{\imath}\colon \Xi \to \bar{\Xi}\colon X \to \bar{X}$ *is an isomorphism.* b) *If* $U = \bar{\imath}^{-1}(\bar{\Xi}^\perp)$ (*and* $V = \bar{\imath}^{-1}(\bar{\Xi}^0)$), *then* $\Xi^\perp \subset U$.

Proof. a) $X \in \Xi$ implies $\bar{X} \in \bar{\Xi}$, since $\bar{X}(0) \in B_{g(0)}$ by the definition of Ξ and since $\bar{X}(\frac{1}{2}) = (X(\frac{1}{2}), X(\frac{1}{2})) \in \Delta_{g(\frac{1}{2})}$. ($\bar{X}$ is restricted to $[0,\frac{1}{2}]$.)

$\bar{\imath}$ is linear; for $\bar{\imath}(aX+Y) = (aX+Y, (aX+Y)^-) = (aX+Y, aX^- + Y^-) = a(X,X^-) + (Y,Y^-) = a\bar{\imath}(X) + \bar{\imath}(Y)$. $\bar{\imath}$ is injective; for $\bar{\imath}(X) = 0$ implies $(X, X^-) = 0$ on $[0,\frac{1}{2}]$. Hence $X = 0$ and $X^- = 0$ on $[0,\frac{1}{2}]$, or $X = 0$ on $[0,1]$. $\bar{\imath}$ is surjective; for $\mathfrak{X} = (X, Y^-) \in \bar{\Xi}$ implies X and Y are vectorfields along parts of σ, defined on the intervals $[0,\frac{1}{2}]$ and $[\frac{1}{2},1]$ respectively. Since $\mathfrak{X}(\frac{1}{2}) \in \Delta_{g(\frac{1}{2})}$ they can be pieced together to a broken C^∞ vectorfield Z along σ, and it is trivial that $\bar{\imath}(Z) = \mathfrak{X}$. b) If $X \in \Xi^\perp$, then $\langle X, \sigma_*\rangle = 0$. Hence $\langle \bar{X}, g_*\rangle = \langle X, \sigma_*\rangle - \langle X^-, \sigma_*^-\rangle = 0$. Thus $\bar{\imath}(X) = \bar{X} \in \bar{\Xi}^\perp$, proving $X \in U$.

Remark. U is strictly bigger than $\Xi^\perp$. It contains also all vectorfields along σ of the form $\psi\sigma_*$, where ψ is a broken C^∞ function satisfying (1) $\psi = \psi^-$ and (2) $\psi(0)(\sigma_*(0), \sigma_*(1)) \in B_{g(0)}$. (In case $B = \Delta$ and σ is periodic, (2) represents no additional restriction. In case $B = S \times T$, (2) implies $\psi(0) = 0$.) V consists of all vectorfields X along σ such that $X = \phi\sigma_*$, where $\phi = -\phi^-$ and $\phi(0) = 0 = \phi(1)$. For $X \in V$ if and only if $\bar{X} = \beta g_*$ where $\beta(0) = \beta(\frac{1}{2}) = 0$. But $\bar{X} = (X, X^-)$ and $g_* = (\sigma_*, -\sigma_*^-)$. Hence $X = \beta\sigma_*$

and $X^- = -\beta\sigma_*^-$ on $[0,\frac{1}{2}]$, or $X = -\beta^-\sigma_*$ on $[\frac{1}{2},1]$. Put $\phi(t) = \beta(t)$ if $t \in [0,\frac{1}{2}]$, $= -\beta^-(t)$ if $t \in [\frac{1}{2},1]$. Then $\phi = -\phi^-$ and $0 = \phi(0) = -\phi(1)$.

LEMMA 2.5. a) *On $U \times U$ $\sqrt{2}\bar{\imath}^*(I_{B\Delta}) = I_B + I$, where I is the positive semi-definite form on $\Xi \times \Xi$ defined by $I(X,Y) = (1/k)\int_0^1 \langle X'^0, Y'^0\rangle$ and X^0 is the component of X tangential to σ_*.*

b) *The subspace V of Ξ is totally singular relative to both $\bar{\imath}^*(I_{B\Delta})$ and I_B.*

Note. $X^{0\prime} = X'^0$, and so $I(X,X) = 0$ if and only if $X^{0\prime} = 0$ if and only if $X^0 = A\sigma_*$, A constant.

Proof of a). Suppose $X, Y \in U$. Then

$$\begin{aligned}
k\sqrt{2}\bar{\imath}^*(I_{B\Delta})(X,Y) &= k\sqrt{2}I_{B\Delta}(\bar{X},\bar{Y}) = \int_0^{\frac{1}{2}} (\langle \bar{X}', \bar{Y}'\rangle - \langle \bar{R}_{g_*\bar{X}}g_*, \bar{Y}\rangle) \\
&\qquad + \text{End term} \\
&= \int_0^{\frac{1}{2}} (\langle X', Y'\rangle + \langle -X'^-, -Y'^-\rangle - \langle R_{\sigma_* X}\sigma_* Y\rangle - \langle R_{-\sigma_*^- X^-} - \sigma_*^-, Y^-\rangle) \\
&\qquad + \text{End term} \\
&= \int_0^{\frac{1}{2}} (\langle X', Y'\rangle - \langle R_{\sigma_* X}\sigma_*, Y\rangle) - \int_1^{\frac{1}{2}} (\langle X', Y'\rangle - \langle R_{\sigma_* X}\sigma_*, Y\rangle) \\
&\qquad + \text{End term} \\
&= \int_0^1 (\langle X', Y'\rangle - \langle R_{\sigma_* X}\sigma_*, Y\rangle) + \text{End term} \\
&= \int_0^1 (\langle X'^{\perp}, Y'^{\perp}\rangle - \langle R_{\sigma_* X^{\perp}}\sigma_*, Y^{\perp}\rangle) + \int_0^1 \langle X'^0, Y'^0\rangle + \text{End term} \\
&= kI_B(X,Y) + kI(X,Y).
\end{aligned}$$

Proof of b). If $X \in V$, $Y \in \Xi$, then $0 = I_{B\Delta}(\bar{X},\bar{Y}) = \bar{\imath}^*(I_{B\Delta}(X,Y)$ because $\bar{X} \in \bar{\Xi}^0$ and $\bar{\Xi}^0$ is a totally singular subspace of $\bar{\Xi}$ relative to $I_{B\Delta}$. On the other hand we showed previously that $X \in V$ implies $X = \phi\sigma_*$, where $\phi = -\phi^-$ and $\phi(0) = 0 = \phi(1)$. Hence

$$I_B(X,Y) = (1/k)\int_0^1 \Gamma(X^{\perp}, Y^{\perp}) + (1/k)\langle S^B{}_{g_*(0)}\bar{X}(0), \bar{Y}(0)\rangle = 0$$

This proves b).

Lemma 2.5 shows that for a comparison of I_B and $I_{B\Delta}$ (via $\bar{\imath}^*$) we can restrict attention to the space U. Define $P = \sqrt{2}\bar{\imath}^*(I_{B\Delta}) \mid U \times U$. Because $\bar{\imath}$ is an isomorphism of U with $\bar{\Xi}^{\perp}$, the index and augmented index of P equal those of $I_{B\Delta} \mid \bar{\Xi}^{\perp} \times \bar{\Xi}^{\perp}$. While $\bar{\Xi}^{\perp}$ is the "right" space for the boundary

problem $(B \times \Delta, g)$ in the sense that the index and nullity of $I_{B\Delta} \mid \bar{\Xi}^{\perp} \times \bar{\Xi}^{\perp}$ give the relevant geometric information regarding shorter paths in $\bar{\Omega}(B, \Delta)$ near g, U (or Ξ) is not "right" for the problem (B, σ) because the nullity $n(I_B \mid U \times U)$ is not finite. The appropriate space to consider is W, defined below.

$$
\begin{aligned}
W &= \{X \in U \mid X = X^{\perp} + a\sigma_*, \text{ where } a \text{ is a constant}\} \\
&= \{X \in U \mid X^{0\prime} = 0\} \qquad (2.3)\\
U^0 &= \{X \in U \mid X \parallel \sigma_* \text{ and } X(0) = 0 = X(1)\}
\end{aligned}
$$

(Recall that $U = \bar{\imath}^{-1}(\bar{\Xi}^{\perp})$.)

Remark. If $B = S \times T$, then $W = \{X \mid (X(0), X(1)) \in S_{\sigma(0)} \times T_{\sigma(1)}$ and $X = X^{\perp}\}$. If $B = \Delta$ and σ is periodic, then W includes the subspace $\{\sigma_*\}$, and equals $\Xi^{\perp} + \{\sigma_*\}$ where $\Xi^{\perp}$ is the space of all broken C^∞ closed vector-fields $\perp \sigma_*$. In fact, this sum is orthogonal with respect to the form I_Δ, and I_Δ restricted to $\{\sigma_*\} \times \{\sigma_*\}$ is 0.

LEMMA 2.6. *$U = U^0 \underset{P}{\oplus} W$, where $\underset{P}{\oplus}$ means direct orthogonal sum with respect to P. Moreover,*

$$P \mid U^0 \times U^0 = I \mid U^0 \times U^0 \text{ and } P \mid W \times W = I_B \mid W \times W.$$

Proof. We first show that this sum is direct. If $X \in U$, then $X = X^{\perp} + \phi\sigma_*$ where $\phi = k^{-2}\langle X, \sigma_*\rangle$, and $\langle \bar{X}, g_*\rangle = 0$. But $\langle \bar{X}, g_*\rangle = \langle X, \sigma_*\rangle - \langle X^{-}, \sigma_*^{-}\rangle = \langle \phi\sigma_*, \sigma_*\rangle - \langle \phi^{-}\sigma_*^{-}, \sigma_*^{-}\rangle = (\phi - \phi^{-})k^2$. Hence $\phi = \phi^{-}$. $X = (X^{\perp} + \phi(0)\sigma_*) + (\phi - \phi(0))\sigma_*$, and since $\phi(0) = \phi(1)$, $Y = (\phi - \phi(0))\sigma_*$ is 0 at 0 and 1. So $\bar{Y}(0) = 0 \in B_{g(0)}$. Further

$$
\begin{aligned}
\langle \bar{Y}, g_*\rangle &= \langle (\phi - \phi(0)\sigma_*, \sigma_*\rangle - \langle (\phi - \phi(0))^{-}\sigma_*^{-}, \sigma_*^{-}\rangle \\
&= (\phi - \phi(0))k^2 - (\phi^{-} - \phi(0))k^2 = 0.
\end{aligned}
$$

Thus $Y \in U^0$. It follows that $X^{\perp} + \phi(0)\sigma_* = X - Y \in U$ and hence in W. Because $U^0 \cap W = \{0\}$, we have proved that $U = U^0 + W$.

Next, suppose $X \in U^0$, $Y \in W$.

$$
\begin{aligned}
P(X, Y) = I_B(X, Y) + I(X, Y) &= (1/k)\int_0^1 \Gamma(X^{\perp}, Y^{\perp}) \\
&+ (1/k)\langle S^B{}_{g_*(0)}\bar{X}(0), \bar{Y}(0)\rangle + (1/k)\int_0^1 \langle X'^0, Y'^0\rangle = 0,
\end{aligned}
$$

because $X^{\perp} = 0$, $\bar{X}(0) = 0$ and $Y^{0\prime} = 0$. The same formula shows the last part of Lemma 2.6, if we let both $X, Y \in U^0$ or both $X, Y \in W$.

LEMMA 2.7. $i(P) = i(I_B \mid W \times W) = i(I_B)$, $n(P) = n(I_B \mid W \times W)$.

Here $i(P)$ = index of P, $n(P)$ = nullity of P. (At this stage it has not been shown that any of these numbers are finite.)

Proof. $i(I_B) = i(I_B \mid U \times U)$ is a consequence of Lemma 2.5b). In the rest of this proof I_B will be understood to mean $I_B \mid U \times U$. Thus $P = I_B + I$, where I is positive semi-definite. In fact, $I \mid U^0 \times U^0$ is positive definite. (For we have shown before that $I(X,X) = 0$ if and only if $X^0 = A\sigma_*$. If also $X \in U^0$, then $X = X^0$ and $A = 0$.) Because $P = I_B + I$ and I is positive semi-definite, $i(P) \leqq i(I_B)$. Now suppose $I_B < 0$ on N (meaning $I_B \mid N \times N$ is negative definite), where N is a subspace of U. Let $N_W = \pi_W(N)$, where π_W is the P-orthogonal projection of U on W. $\ker(\pi_W \mid N) \subset U^0$. Hence, if $X \in \ker(\pi_W \mid N)$, then $X \in U^0$ and $I_B(X,X) = 0$, contradicting the fact that $I_B < 0$ on N, unless $X = 0$. Hence N_W is isomorphic to N and $I_B < 0$ on N_W since $I_B(X,X) = I_B(\pi_W X, \pi_W X)$. Thus, if N is a maximal subspace of U on which $I_B < 0$, there always exists a subspace N_W of W, isomorphic with N, on which $I_B < 0$. Since $P \mid N_W \times N_W = I_B \mid N_W \times N_W$, it follows that $i(P) \geqq i(I_B)$. Combined with the above inequality this yields the first part of Lemma 2.7.

Let K be a maximal totally singular subspace of U relative to P. Then $K \subset W$. For if $X \in K$, $X = X_{U^0} + X_W$ and $P(X,Y) = 0$ for all $Y \in U$. In particular, $0 = P(X, X_{U^0}) = P(X_{U^0}, X_{U^0}) = I(X_{U^0}, X_{U^0})$. Hence $X_{U^0} = 0$, showing that $X = X_W \in W$, i.e. $K \subset W$. Since $P \mid W \times W = I_B \mid W \times W$, we obtain $n(P) = n(I_B \mid W \times W)$, which concludes the proof of Lemma 2.7.

Summarising, we have proved

THEOREM 2.1. *Let I_B and $I_{B\Delta}$ be the index forms defined on $\Xi \times \Xi$ and $\hat{\Xi} \times \hat{\Xi}$ by (2.1) and (2.2) respectively, and let W be the subspace of Ξ consisting of all broken C^∞ vectorfields X along σ, satisfying $X^{0\prime} = 0$ and $(X(0), X(1)) \in B_{(\sigma(0),\sigma(1))}$, where X^0 is the component of X parallel to σ_*. Then the following relations hold between their indices and nullities*

$$i(I_{B\Delta}) = i(I_B \mid W \times W) = i(I_B)$$
$$n(I_{B\Delta} \mid \hat{\Xi}^\perp \times \hat{\Xi}^\perp) = n(I_B \mid W \times W).$$

As a corollary of this theorem we shall obtain a symmetric version of the Ambrose index theorem [2], as well as a generalization of it. With B, Δ and g as above, the Ambrose index theorem states that the index of $I_{B\Delta}$ on $\hat{\Xi}^\perp \times \hat{\Xi}^\perp$ is the sum of the orders of the conjugate points of the ordered pair (Δ, B) on g plus a convexity term $c_{\Delta B}$. The latter measures the "relative convexity" of Δ and a nearby $B(t)$ where $B(t)$ is the manifold B (locally)

"pushed over" toward Δ along g via Jacobifields (see also Section 4); precisely $c_{\Delta B}$ = index of $I_{B(t)\Delta}$ when t is sufficiently close to $\frac{1}{2}$. But by Theorem 2.1 applied to the interval $[t, 1-t]$ $i(I_{B(t)\Delta}) = i(I_{B(t)})$ where $I_{B(t)}$ is the index form for the problem $(B(t), \sigma \mid [t, 1-t])$. We write c_B for $c_{\Delta B}$.

Definition. t is a *symmetric conjugate point* for B on σ, if $g(t) = (\sigma(t), \sigma(1-t))$ is a conjugate point of the ordered pair (Δ, B) in the sense of Ambrose [2].

Thus in particular $t < \frac{1}{2}$. Explicitly, t is a symmetric conjugate point for B on σ if there exist Jacobifields X, Y and Z along σ, not all equal, such that if $\mathfrak{X} = (X, Y^-)$ then $\mathfrak{X} \in \bar{\Xi}^\perp$, $\mathfrak{X}'(0) = S^B{}_{g_*(0)}\mathfrak{X}(0)$, $X(t) = Z(t)$, $Y(1-t) = Z(1-t)$, and

$$(X'(t) - Z'(t), -Y'(1-t) + Z'(1-t)) \perp B_1(t)$$

where $B_1(t) = \{(u, v) \in (M \times M)_{g(t)} \mid u = U(t),\ v = V(t)$ for Jacobifields U, V along σ satisfying $(U(0), V(1)) \in B_{g(0)}$ and $(U'(0), -V'(1)) = S^B{}_{g_*(0)}(U(0), V(1))\}$.

Theorem 2.1 in conjunction with the Ambrose theorem yields

THEOREM 2.2. *Let B be a submanifold of $M \times M$, σ a geodesic in M defined on the interval $[0,1]$ such that the geodesic $g = (\sigma, \sigma^-)$ hits B perpendicularly at $g(0)$, and let W be the linear space of all broken C^∞ vectorfields X along σ such that $X^{0\prime} = 0$ and $(X(0), X(1)) \in B_{g(0)}$, where X^0 is the component of X parallel to σ_*. If $i(B)$ is the index of $I_B \mid W \times W$, where I_B is defined on $\Xi \times \Xi$ by (2.1), then*

$$i(I_B) = i(B) = \sum_{0<t<\frac{1}{2}} n(t, B) + c_B < \infty$$

where $n(t, B)$ is the order of t as a symmetric conjugate point for B on σ and c_B is the convexity term referred to above.

Remarks. 1. If $B = S \times T$ this theorem gives a symmetric version of the Ambrose theorem, in which both submanifolds S and T are pushed toward each other at the same speed. In particular, $c_{S\times T}$ is the index of $I_{S(t)T(1-t)}$ when t is sufficiently near $\frac{1}{2}$. We describe a special case of some interest. Let S be arbitrary, but T only the point $\sigma(1)$. Then t is a symmetric conjugate point of $S \times T$ if $1-t$ is a focal point of the boundary condition $S(t)$, that is if there exist Jacobifields X, Y along σ, not both trivial, such that $X(0) \in S_{\sigma(0)}$, $X'(0) = S^S{}_{\sigma_*(0)}(X(0))$, $X(t) = Y(t)$, $Y(1-t) = 0$ and

$X'(t) - Y'(t) \perp S_1(t)$. If we assume known that the focal points of a submanifold along a given geodesic perpendicular to it cannot occur arbitrarily close to the submanifold (see Section 3), and if we furthermore assume that $\frac{1}{2}$ is not a strong focal point of S (see [2], or Section 4), then the continuous dependence of $S(t)$ on t near $\frac{1}{2}$ implies (by the continuity theorems of Section 4) that $c_{S\times T} = 0$. If S also reduces to a point, so that $S \times T$ is just the point $(\sigma(0), \sigma(1))$, the symmetric conjugate points are just symmetric pairs of conjugate points in the usual sense, i.e. pairs $(t, 1-t)$ for which there exist non-trivial Jacobifields vanishing both at t and at $1-t$. In this case $c_{S\times T} = 0$, trivially. We summarise this remark in a

COROLLARY. *If $B = S \times \{\sigma(1)\}$, and if $\frac{1}{2}$ is not a strong focal point of S, then $i(B) = \sum_{0<t<\frac{1}{2}} n(t, B)$.*

(Note that in the simple case when $S = \{\sigma(0)\}$ this corollary is a consequence of the continuity Theorem 4.3 of Morse.)

2. *If $B = \Delta$ and σ is periodic, a new description of the index of closed geodesics results.*

3. In order to obtain Theorem 2.2 we pushed the manifold B toward Δ. Another theorem would result if we pushed Δ toward B (though in the periodic case when $B = \Delta$, these two theorems would be no different).

3. Index theorems of Morse. This section is devoted to a proof of the general index theorem (3.2) as it was apparently first indicated by Morse [13]. This important index theorem does not seem to be widely known, perhaps because Morse's results are phrased (in great generality) in the language of the classical calculus of variations. Another proof of it is given by Edwards [7]. This index theorem of Morse does not seem to imply a recent one of Ambrose [2], nor conversely.

The geometric setting is as in Section 2. We consider the space W of all broken C^∞ vectorfields X along the geodesic σ on the interval $[0,1]$ such that $(X(0), X(1)) \in B_{(\sigma(0),\sigma(1))}$ and $X^{0\prime} = 0$, where X^0 is the component of X parallel to σ_*. We assume the problem is critical (i.e. the lift g of σ is perpendicular to B at $g(0)$). The index form I_B is defined as in (2.1) on $W \times W$. We seek to determine the index and nullity of I_B in terms of the conjugate points of $\sigma(0)$ along σ and the index and nullity of I_B restricted to a finite dimensional space of Jacobifields along σ.

We shall make use of the well-known focal point theorem, and state it explicitly. Proofs can be found in [2], [9], [12], [14].

THEOREM 3.1. *Let W be the space of all broken C^∞ vectorfields along and perpendicular to the geodesic σ, tangent to the submanifold S of M, (which is assumed to hit σ perpendicularly at $\sigma(0)$) and vanishing at $\sigma(1)$. Let I_{S0} $(=I_{S\times\sigma(1)})$ be the index form of the second variation for this problem, restricted to $W\times W$, and $f_S(t)$ the order of $\sigma(t)$ as a focal point of S along σ. Then*

$$\text{index of } I_{S0}=i(I_{S0})=\sum_{0<t<1} f_S(t)<\infty$$

$$\text{nullity of } I_{S0}=n(I_{S0})=f_S(1).$$

For the proof of Theorem 3.2 only the special case of Theorem 3.1 where S is the point $\sigma(0)$, is needed. But the same methods that lead to Theorem 3.2 can be used to obtain the index of I_B in terms of the focal points of B along g and the index of I_B restricted to a finite dimensional space of Jacobi-fields along σ with breaks at one point, and for this Theorem 3.1 in full strength is needed.

We now develop the algebra needed to prove Theorems 3.2 and 3.3. In the most usual ("non-degenerate") case, when $\sigma(0)$ is not conjugate to $\sigma(1)$, the proof of Theorem 3.2 is so exceedingly simple that we believe a separate proof for this case is justified. Lemma 3.1 below deals with the complications that arise in the degenerate case (which are often algebraic rather than geometric in nature). Since this lemma does not seem to generalise to the cases where V and W are not finite dimensional, it is necessary to reduce the problem geometrically to a finite dimensional case, to which Lemma 3.1 applies.

Notation. Infinite (∞) means "not finite," and $<\infty$ means "finite." V is a vectorspace over the reals, $\partial(V)=$ dimension of V. $f\colon V\times V\to R$ is a bilinear symmetric form, $p=f\mid W\times W$ is the restriction of f to $W\times W$. $f<0$ on a subspace T means $f\mid T\times T$ is negative definite.

$$W^\perp=\{x\in V\mid f(x,y)=0 \text{ for all } y\in W\},$$
$$S_f=\{x\in V\mid f(x,y)=0 \text{ for all } y\in V\}.$$

$q=f\mid W^\perp\times W^\perp$. $\nu(f,p)=\partial(W\cap S_f)$ and $i(f)\,(a(f),n(f))$ is the index (augmented index, nullity) of f.

LEMMA 3.1. *Assume $\partial(V)<\infty$. Then* $i(f)=a(p)+i(q)-\nu(f,p)$ $=i(p)+n(p)+i(q)-\nu(f,p)$.

Proof. $W\cap S_f$ is a subspace of S_p. Let U be any compliment of S_p in W while $\bar{U}$ is any compliment of $W\cap S_f$ in S_p. Then

$$W = U \oplus \bar{U} \oplus (W \cap S_f) = U \oplus S_p,$$

where $\oplus$ means orthogonal sum relative to f (or any form induced by f in appropriate subspaces.) Thus $i(p) = i(p \mid U \times U)$. Notice that $W^{\perp} \cap W = S_p \subset S_q$, and that $W^{\perp} = T \oplus S_q = T \oplus \bar{T} \oplus S_f$, where T is any compliment of S_q in $W^{\perp}$. Put $k = \partial(U)$. P is called an f-hyperbolic plane if $\partial(P) = 2$ and if $i(f \mid P \times P) = 1$.

Lemma 3.1A. *There exist f-hyperbolic planes $P_1 \cdots P_k$ such that $V = U \oplus P_1 \oplus \cdots \oplus P_k \oplus T \oplus S_f$.*

Proof. We refer to Theorem 3.8 of Artin [3]. The algebra involved is standard and forms the main step in the proof of Witt's theorem.

From Lemma 3.1A we now deduce Lemma 3.1 in the following manner. $i(f \mid U \times U) = i(p \mid U \times U) = i(p)$, $i(f \mid T \times T) = i(q \mid T \times T) = i(q)$, *and* $i(f \mid P_i \times P_i) = 1$. By Lemma 3.1A $i(f) = i(p) + k + i(q) = i(p) + k + \nu(f,p) - \nu(f,p) + i(q) = a(p) - \nu(f,p) + i(q)$, which proves Lemma 3.1.

Lemma 3.2. *Suppose $V = W \oplus U$, where $\partial(U) < \infty$ (but V and W may not be finite dimensional) and $\oplus$ denotes orthogonal direct sum with respect to f. Suppose $i(p) < \infty$, and let $q = f \mid U \times U$. Then $i(f) = i(p) + i(q)$.*

Proof. $i(f) \geqq i(p) + i(q)$ is obvious. Before showing the opposite inequality, we eliminate the possibility that $i(f) = \infty$. For if $i(f) = \infty$, there exist subspaces D_m of V of arbitrarily large dimension m on which $f < 0$. Then D_m is contained in the finite dimensional subspace $K_m = \pi_W D_m \oplus U$ of V, where π_W denotes the orthogonal projection on W. Since the lemma is true for finite dimensional spaces (then a special case of Lemma 3.1), we obtain $i(p) + i(q) \geqq i(p \mid \pi_W D_m \times \pi_W D_m) + i(q) = i(f \mid K_m \times K_m) \geqq m$. Since $i(p)$ and $i(q)$ are fixed (independent of m) and finite, this would lead to a contradiction if m is unbounded.

Thus, letting D be a maximal finite dimensional space on which $f < 0$, the same reasoning as above shows $i(p) + i(q) \geqq i(f \mid K \times K) = i(f)$, completing the proof of Lemma 3.2.

We can now turn to the proof of the general index theorem of Morse. Let $J_{B_1} = \{X \in W \mid X$ is unbroken Jacobifield along $\sigma\}$, and $J_B{}^* = \{X \in J_{B_1} \mid X$ satisfies the boundary condition $B\}$. Finally, put $W^0 = \{X \in W \mid X(0) = 0 = X(1)\}$, $I^0 = I_B \mid W^0 \times W^0$ and $E_B = I_B \mid J_{B_1} \times J_{B_1}$. From the integration by parts formula

$$(3.1)\qquad kI_B(X,Y)=\int_0^1\langle -R_{\sigma_* X\sigma_*}-X'',Y\rangle+\sum\langle X'(t_i^-)-X'(t_i^+),Y(t_i)\rangle \\ +\langle -\bar{X}'(0)+S^B{}_{g_*(0)}(\bar{X}(0)),\bar{Y}(0)\rangle$$

applied to $X, Y \in W$ (we have used the property that $X^{0\prime}=0$ to obtain this formula) we conclude that $E_B(X,Y)=(1/k)\langle -\bar{X}'(0)+S^B{}_{g_*(0)}(\bar{X}(0)),\bar{Y}(0)\rangle$ if $X, Y \in J_{B_1}$.

Let $n(t)$ denote the order of $\sigma(t)$ as a conjugate point of $\sigma(0)$ along σ (in the ordinary sense). $i(I_B)$ is the index of I_B on $W\times W$ and $i(E_B)$ the index of E_B on $J_{B_1}\times J_{B_1}$.

THEOREM 3.2. (Morse)

a) $i(I_B)=\sum_{0<t\leqq 1} n(t)+i(E_B)-\nu(B)<\infty$

b) $n(I_B)=$ dimension $J_B{}^*$.

Here $\nu(B)=\partial(J_B{}^*\cap J^0)=$ dimension of the linear space of all Jacobifields in W satisfying B and vanishing at $\sigma(0)$ and $\sigma(1)$.

Note. $\nu(B)=0$ if $\sigma(0)$ is not conjugate to $\sigma(1)$ along σ; (the non-degenerate case).

I. *Proof of Theorem* 3.2 *in the non-degenerate case.* We claim $W=W^0\oplus J_{B_1}$ where $\oplus$ denotes the orthogonal direct sum with respect to the index form I_B. If our claim is correct, Theorem 3.2 a) is immediate by Lemma 3.2 and Theorem 3.1.

We prove the claim by observing that if $X\in W$, there is a unique Jacobifield $Y\in W$ such that $\bar{Y}(0)=\bar{X}(0)$, because $\sigma(0)$ is not conjugate to $\sigma(1)$ and $X\in W$ implies $X^{0\prime}=0$. Thus $X=(X-Y)+Y$ uniquely, where $X-Y\in W^0$ and $Y\in J_{B_1}$. It remains to show this sum is orthogonal with respect to I_B. Let $Y\in W^0$, $X\in J_{B_1}$. Then (3.1) shows that $I_B(X,Y)=0$. Hence our claim has been proved. Theorem 3.2 b) is proved as below.

II. *Proof of Theorem* 3.2. Part b). Suppose $I_B(X,Y)=0$ for all $Y\in W$. In the usual way by making appropriate choices for Y, it follows from (3.1) that X is smooth, that X is Jacobi and that X satisfies B. Hence the nullspace of I_B is contained in $J_B{}^*$, and the converse is trivial by (3.1). Part a). As in Section 2 we "lift" the problem to $M\times M$. Choose a normal sequence $\{t_i\}$ for $[0,\frac{1}{2}]$ on $g=(\sigma,\sigma^-)$ ([2], page 64, or Section 4.). Let $\bar{\mathfrak{B}}$ be the corresponding space of broken Jacobifields perpendicular to g with breaks at the t_i, starting tangent to B and ending tangent to Δ. $\bar{W}$, by

definition, is isomorphic to W via i, in fact conformally relative to the forms $I_{B\Delta}$ and I_B. There is a projection map $\bar{\beta}: \bar{\Xi}^{\perp} \to \bar{\mathcal{B}}$, and as in Theorem 1 of [2], we obtain that $i(I_{B\Delta} \mid \bar{W} \times \bar{W}) = i(I_{B\Delta} \mid \bar{W} \cap \bar{\mathcal{B}} \times \bar{W} \cap \bar{\mathcal{B}})$. Defining $\mathcal{B} = \bar{i}^{-1}(\bar{W} \cap \bar{\mathcal{B}})$, we consequently obtain that $i(I_B) = i(I_B \mid \mathcal{B} \times \mathcal{B})$. $\mathcal{B}$ is described explicitly as the space of all broken Jacobifields X along σ with breaks at the t_i and $1 - t_i$, such that $X^0 = A\sigma_*$ where A is constant, and such that $\bar{X}(0) \in B_{g(0)}$.

We apply Lemma 3.1 to $\mathcal{B}$, with $W = W^0 \cap \mathcal{B}$ and $f = I_B \mid \mathcal{B} \times \mathcal{B}$, and claim (1) $i(f) = i(I_B)$, (2) $a(p) = a(I^0 \mid W^0 \times W^0)$, (3) $i(q) = i(E_B)$, (4) $\nu(f, p) = \nu(B)$. Interpreting the right side of (2) in the light of Theorem 3.1 in the "fixed ends" case, i.e. $S = \{\sigma(0)\}$, we obtain Theorem 3.2 a).

It remains to verify (2)-(4). Proof of (2). Since $\{t_i, 1 - t_i\}$ is a normal sequence for $[0,1]$ on σ, (2) is a consequence of the reduction described in Theorem 1 of [2], as applied to the fixed ends problem. Proof of (3). We must show that the perpendicular space to $W^0 \cap \mathcal{B}$ relative to I_B in $\mathcal{B}$ is J_{B_1}. So suppose $X \in (W^0 \cap \mathcal{B})^{\perp}$. Since X is broken Jacobi in W, formula (3.1) gives that

$$\begin{aligned} 0 = kI_B(X, Y) &= \sum \langle X'(s_j^-) - X'(s_j^+), Y(s_j) \rangle + kE_B(X, Y) \\ &= \sum \langle X'(s_j^-) - X'(s_j^+), Y(s_j) \rangle \text{ for all } Y \in W^0 \cap \mathcal{B}. \end{aligned}$$

(As usual the summation is over the breaks of X.) From appropriate choices of Y, namely such that $Y(s_j) = 0$ except when $j = r$, we deduce that $X'(s_r^-) - X'(s_r^+)$ is tangential to $\sigma_*(s_r)$. But the tangential component of X' is 0, hence X has no break at s_r. Thus $X \in J_{B_1}$ and $(W^0 \cap \mathcal{B})^{\perp} \subset J_{B_1}$. The converse is immediate by (3.1). Proof of (4). Observe the definition of $\nu(f, p)$ and use b) above.

We next prove the analog of Theorem 3.2 with focal points of B rather than conjugate point of $\sigma(0)$.

THEOREM 3.3. (Morse)

$$i(I_B) = \sum_{0 < t \leq \frac{1}{2}} f_B(t) + i(F_B) - \nu(B, \tfrac{1}{2}) < \infty.$$

Here $f_B(t)$ = *order of* $g(t)$ *as a focal point of* B *along* $g = (\sigma, \sigma^-)$, $F_B = I_B \mid J(B, \frac{1}{2}) \times J(B, \frac{1}{2})$, $J(B, \frac{1}{2}) = \{X \in W \mid X$ *satisfies* B *and is broken Jacobi with at most one break at* $\sigma(\frac{1}{2})\}$, *and* $\nu(B, \frac{1}{2}) = \partial\{X \in W \mid X$ *is smooth Jacobi satisfying* B *and* $X(\frac{1}{2}) = 0\}$.

COROLLARY. *If* $B = S \times T$, *then*

$$i(I_{ST}) = \sum_{0<t\leqq\frac{1}{2}} f_S(t) + \sum_{0<t\leqq\frac{1}{2}} f_T(t) + i(F_{ST}) - \nu(ST, \tfrac{1}{2}) < \infty$$

Remark. This corollary appears in a slightly greater generality in Morse [13], the division of σ occurring at an arbitrary point, not necessarily the midpoint. The method of proof is precisely the same, except that in this case of separate end manifolds there is no need to consider the problem in $M \times M$.

If Theorem 3.3 is applied to the case when $B = \Delta$ and σ is periodic, no new statement results; we only obtain a restatement of Theorem 3.2 for $B = \Delta$.

Proof of Theorem 3.3. We use the same notation as above. In addition $\bar{\Xi}^{\perp}(B, 0) = \{\mathfrak{X} \mid \mathfrak{X}$ is broken C^∞ along and perpendicular to g, starting tangent to B and ending with $\mathfrak{X}(\frac{1}{2}) = 0\}$. I_{B0} is the index form for the problem $(B \times \{g(\frac{1}{2})\}, g)$ in $M \times M$. Recall that $\bar{W} = \bar{\imath}(W)$. Then the augmented index $a(I_{B0}) = a(I_{B0} \mid \bar{\Xi}^{\perp}(B,0) \cap \bar{W} \times \bar{\Xi}^{\perp}(B,0) \cap \bar{W})$. Next choose a normal sequence for $[0, \frac{1}{2}]$ on g, and let $\bar{\mathfrak{B}}$ be the associated space of broken Jacobifields, starting tangent to B and ending tangent to Δ. Denote $\bar{\Xi}^{\perp}(B, 0) \cap \bar{W} \cap \bar{\mathfrak{B}}$ by $\bar{\mathfrak{F}}$. Then, as in the previous proof, we obtain

$$i(I_{B\Delta} \mid \bar{\mathfrak{B}} \cap \bar{W} \times \bar{\mathfrak{B}} \cap \bar{W}) = i(I_{B\Delta}) = i(I_B),\ a(I_{B0} \mid \bar{\mathfrak{F}} \times \bar{\mathfrak{F}}) = a(I_{B0}).$$

Hence, by Lemma 3.1, $i(I_B) = a(I_{B\Delta} \mid \bar{\mathfrak{F}} \times \bar{\mathfrak{F}}) + i(I_{B\Delta} \mid \bar{\mathfrak{F}}^{\perp} \times \bar{\mathfrak{F}}^{\perp}) - \nu$. This implies Theorem 3.3 when we notice that $I_{B\Delta} \mid \bar{\mathfrak{F}} \times \bar{\mathfrak{F}} = I_{B0} \mid \bar{\mathfrak{F}} \times \bar{\mathfrak{F}}$ and apply Theorem 3.1 to express the index of I_{B0} in terms of the focal points of B along g; when we observe that $\nu = \nu(B, \frac{1}{2})$; and when we prove that the orthogonal space to $\bar{\mathfrak{F}}$ in $\bar{W} \cap \bar{\mathfrak{B}}$ relative to $I_{B\Delta}$ is precisely $J(B, \frac{1}{2})$. The latter is seen in the usual manner from the integration by parts formula for $I_{B\Delta}$.

Remark 1. In the case of separate end manifolds, $B = S \times T$, Morse states a slight generalization of Theorem 3.2 which is named "generalized concavity theorem." It says. $0 = b_0 < b_1 < \cdots < b_p < b_{p+1} = 1$, if $k_i =$ the number of conjugate points of b_i (along σ) in $(b_i, b_{i+1}]$ counted with multiplicities, then $i(I_{ST}) = \sum_{i=0}^{p} k_i + i(I_{ST} \mid J_p \times J_p) - \nu_p$ where $J_p = \{X \mid X$ is a broken Jacobifield along σ with breaks at the b_i, everywhere perpendicular to σ, and starting tangent to S, ending tangent to $T\}$. $\nu_p = \partial\{X \mid X$ is unbroken Jacobi, satisfies S and T and vanishes at all the b_i, $i = 0, 1, \cdots, p+1\}$.

The proof is carried through in the same manner as in the case of Theorem 3.2. Let $W_p{}^0 = \{X \mid X$ is broken C^∞ along σ and $X(b_i) = 0$, $i = 0, 1, \cdots, p+1\}$. When Lemma 3.1 is applied to $W_p{}^0$ (or more precisely

to a suitable finite dimensional subspace of broken Jacobifields), and when we observe that $W_p{}^{0\perp}$ relative to I_{ST} is J_p, the statement above follows.

2. A more general theorem of this sort can actually be obtained by the same simple methods. It relates the indices of I_{ST}, I_{SU} and I_{TU} for any three submanifolds S, T and U of M perpendicular to the geodesic σ at $\sigma(0)$, $\sigma(r)$ and $\sigma(1)$, respectively.

4. Continuity properties of the index form.

4.1. Given $B = (\langle,\rangle, S, T, \sigma)$ where $\langle,\rangle$ is a (complete) Riemannian metric on the manifold M, S and T are submanifolds, and σ is a geodesic starting and ending perpendicularly on S and T respectively. Intuitively, we ask: If $i(B)$, the index of the second variation problem associated with the data B, is considered as a function of B, is $i(B)$ stable (under sufficiently small variations in B)? Do the (Ambrose) conjugate points depend continuously on B? It is perhaps no surprise that the answers are affirmative, provided the nullity of the problem B is 0 (Theorems 4.1 and 4.2). But there are two possible topologies on the set of self adjoint boundary conditions, one which corresponds closely to the geometric description of boundary conditions (Section 2), and another (coarser) which is natural when the boundary conditions are regarded as elements of a Grassmann manifold (4.5, Definition II). We shall show also, that if we use the latter topology, the answers to the above questions would in general be no.

The continuity theorems will be formulated analytically due to their generality. The transition from a geometric problem $(\langle,\rangle, S, T, \sigma)$ to a problem in R^d ($d+1=$ dimension of M) is effected by a choice of a parallel-translated frame along σ, the first d vectors of which span the tangent spaces perpendicular to σ_*. Where a change of end-manifolds and Riemannian structure is involved, many such choices will have to be made so as to give continuous or differentiable maps to which the theorems below apply. We remark that these theorems are stated only for separated end conditions, but we proved in Section 2 that the "more general" ones are included as special cases.

4.2. *Notation.* R^d is d-dimensional Euclidean space with the natural inner product $\langle,\rangle$. $K: TR^d \to R^d$ is the usual identification map. Broken C^∞ curves in R^d are sometimes referred to as broken C^∞ vectorfields. If X is such a curve, put $X' = KX_*$ wherever X_* is defined. $\Xi(s,t)$ is the linear space of all broken C^∞ curves defined on $[s,t]$. $\Xi^0(s,t) = \{X \in \Xi(s,t) \mid X(s) = 0 = X(t)\}$. $\mathfrak{M}(n,d)$ is the manifold of linear transformations from R^n to R^d (the additive group of $d \times n$ matrices). $GL(d)$ is the general linear group of automorphisms of R^d. $S(d)$ is the manifold of self-adjoint (sym-

metric) linear transformations of R^d relative to $\langle , \rangle$. A family of curvature transformations depending on a parameter $\alpha \in R^k$ is given. Precisely, $R: R^k \times R^1 \to S(d)$ is a continuous map, such that in addition the maps $R^\alpha = R(\alpha, \cdot): R^1 \to S(d)$ are C^∞ for each $\alpha \in R^k$. (We write R^1 for the reals in order to avoid confusion with the curvature map R). $I_{(\alpha s t)}$ is a bilinear symmetric form defined on $\Xi(s,t) \times \Xi(s,t)$ by

$$(4.1) \qquad I_{(\alpha s t)}(X,Y) = f(\alpha) \int_s^t (\langle X', Y' \rangle - \langle R^\alpha X, Y \rangle)$$

where $f: R^k \to R^1$ is continuous and strictly positive.

4.3. *Jacobifields, associated maps and ordinary conjugate points.* Associated with each α is the Jacobi equation

$$(4.2) \qquad X'' + R^\alpha X = 0.$$

J^α is the $2d$ dimensional solution space of "α-Jacobifields."

$$J^\alpha(s) = \{X \in J^\alpha \mid X(s) = 0\}.$$

For each (α, s, t) define $\Phi^\alpha{}_{st}: R^{2d} \to R^{2d}$ by $\Phi^\alpha{}_{st}(x, x') = (X(t), X'(t))$ where X is the unique α-Jacobifield such that $(X(s), X'(s)) = (x, x')$. Define "evaluation" maps $E^\alpha{}_{st}, E'^\alpha{}_{st}: R^{2d} \to R^d$ and $F^\alpha{}_{st}, F'^\alpha{}_{st}: R^d \to R^d$ by $E^\alpha{}_{st}(x,x') = X(t)$, $E'^\alpha{}_{st}(x,x') = X'(t)$, $F^\alpha{}_{st}(y) = Y(t)$, $F'^\alpha{}_{st}(y) = Y'(t)$ where X is as above, and Y is the unique α-Jacobifield such that $(Y(s), Y'(s)) = (0, y)$.

A skew symmetric bilinear non-degenerate form J is defined on $R^{2d} \times R^{2d}$ by $J: R^{2d} \times R^{2d} \to R^1: ((x,y),(u,v)) \to \langle x, v \rangle - \langle y, u \rangle$ when $x, y, u, v \in R^d$. The set of element of $GL(2d)$ which preserve J form a Lie subgroup H. The following are now consequences of the general theory of differential equations [8].

(I) $\Phi^\alpha{}_{st} \in H$, and $\Phi: R^k \times R^1 \times R^1 \to H: (\alpha, s, t) \to \Phi^\alpha{}_{st}$ is continuous (in fact differentiable for each fixed α) and satisfies $\Phi^\alpha{}_{st} \circ \Phi^\alpha{}_{rs} = \Phi^\alpha{}_{rt}$. (I) implies

$$(\mathrm{II}) \qquad \begin{array}{l} E, E': R^k \times R^1 \times R^1 \to \mathfrak{M}(2d, d): (\alpha, s, t) \to E^\alpha{}_{st}, E'^\alpha{}_{st} \\ F, F': R^k \times R^1 \times R^1 \to \mathfrak{M}(d, d): (\alpha, s, t) \to F^\alpha{}_{st}, F'^\alpha{}_{st} \end{array}$$

are continuous, and differentiable for each fixed α.

As usual we define (s,t) to be an α-conjugate pair (or s α-conjugate to t) if there is a nontrivial $X \in J^\alpha$ such that $X(s) = 0 = X(t)$. The order of (s,t)

is the dimension of the subspace of J^{α} spanned by such X. In particular (s,s) has order d for any α.

LEMMA 4.1. *Let C and D be any compact subsets of R^1 and R^k respectively. Then there exists a positive number δ (depending on C and D in general) such that if $s \in C$, $0 < |s-t| < \delta$, then (s,t) cannot be an α-conjugate pair for any $\alpha \in D$.*

Remarks. 1. For a fixed α this lemma corresponds to the fact that if C is a compact subset of a complete Riemannian manifold M, there exists $\delta > 0$ such that $\exp_m$ is a diffeomorphism on every ball at the origin of M_m of radius less than δ, for any $m \in C$. The parameter α may be regarded as representing a particular Riemannian metric on M (although R^{α} will not in general be the curvature tensor associated with some metric on M).

2. The method of studying the solutions of the differential equation (4.2) via the form (4.1) is classical and applicable to a much larger class of differential equations.

Proof of Lemma 4.1. (a) Given C, D as above. Then there exists $\delta > 0$ such that $I_{(\alpha s t)}$ is positive definite on $J^{\alpha}(s)$ for all $(s,\alpha) \in C \times D$ and $0 < |s-t| < \delta$. From (a) the lemma follows when we observe that if $X \in J^{\alpha}$ and $X(s) = 0 = X(t)$, then $I_{(\alpha s t)}(X,X) = 0$. The claim (*a*) is a consequence of the continuity of the functions F and F' together with the facts that $F^{\alpha}{}_{ss} = 0$ and $F'^{\alpha}{}_{ss} = I_d$ (the identity map) for all (α, s).

4.4. *Normal sequences.* Following [2] we define an α-normal sequence $\{t_i\}$ for $[s,t]$ to be a finite sequence such that $s < t_1 < \cdots < t_p < t$ and such that there are no α-conjugate points of t_i in $[t_{i-1}, t_{i+1}]$ except t_i itself $(i = 1, \cdots, p)$, and no α-conjugate points of s in $(s, t_1]$ and of t in $[t_p, t)$. (The last assumption is actually redundant by the index theorem for fixed endpoints).

LEMMA 4.2. *Given $(\alpha, s, t) \in R^k \times R^1 \times R^1$, $s < t$, and a compact neighborhood D_α of α. Then there exist an α-normal sequence $\{t_i\}$ for $[s,t]$ and neighborhoods U_s and U_t of s and t so that for every $(\bar\alpha, \bar s, \bar t) \in D_\alpha \times U_s \times U_t$ $\{t_i\}$ is an $\bar\alpha$-normal sequence for $[\bar s, \bar t]$.*

Proof. Let C be any compact neighborhood of $[s,t]$. By Lemma 4.1 there is $\delta > 0$ such that $c \in C$ and $0 < |c-d| < \delta$ implies c is not $\bar\alpha$-conjugate to d for all $\bar\alpha \in D_\alpha$. Choose any sequence $\{t_i\}$ such that $s < t_1 < \cdots < t_p < t$ and $\delta/3 < |t_i - t_{i-1}| < \delta/2$ $(i = 1, \cdots, p+1, t_0 = s, t_{p+1} = t)$. Let U_s

$= \{c \mid |c - s| < \delta/3\} \cap C$, $U_t = \{c \mid |c - t| < \delta/3\} \cap C$. With these choices for $\{t_i\}$, U_s and U_t Lemma 4.2 holds.

4.5. *Self adjoint boundary conditions.* We proved (Section 2) that the aspects of a submanifold and a geodesic perpendicular to it relevant to any index problem, are its tangent space at the starting point of the geodesic and its second fundamental form relative to the starting direction of the geodesic. These data are abstracted in

Definition I. [7] A self adjoint boundary condition in dimension $d+1$ is an ordered pair of endomorphisms of R^d, (S_1, S_2), satisfying

$$\text{(B1) } S_1 \text{ is an orthogonal projection,} \qquad \text{(B2) } S_2 \text{ is self adjoint,}$$
$$\text{(B3) } S_2(I_d - S_1) = I_d - S_1.$$

It follows that $S_1S_2 = S_2S_1$. (The dimension $d+1$ refers to the dimension of the space in which the submanifold is imbedded and not to the dimension of the tangent space of the submanifold. The dimension of the latter is the dimension of the image of S_1.) $\mathcal{S} = \mathcal{S}(d)$ is the subset of $S(d) \times S(d)$ consisting of all self adjoint boundary conditions in dimension $d+1$, given the induced topology. As in [2] we also denote the image of the projection S_1 by S_1. Suppose $\partial(S_1) = k$. Associated with (S_1, S_2) is a k-plane p in R^{2d}, which we call the *essence* of (S_1, S_2), defined by $p = \{(S_1x, S_2S_1x) \mid x \in R^d\}$ $= \{(y, S_2y) \mid y \in S_1\}$.

Let $J: R^{2d} \to R^{2d}$ be the bilinear skew symmetric form defined in 4.3. A maximal subspace P of R^{2d} such that $J \mid P \times P = 0$ is called a maximal nullspace of J and has dimension d.

Definition II. [7] A self adjoint boundary condition in dimension $d+1$ is a maximal nullspace of J in R^{2d}.

$\mathcal{B} = \mathcal{B}(d)$ is the subset of the Grassmann manifold $G_{2d,d}$ consisting of all self adjoint boundary conditions in dimension $d+1$ in the sense (II). Let $\mathcal{B}_{2d,k} = \{k\text{-planes } p \text{ in } R^{2d} \mid J \mid p \times p = 0\}$. Thus $k \leqq d$ and $\mathcal{B}_{2d,d} = \mathcal{B}$. $\mathcal{B}^0_{2d,k} = \{p \in \mathcal{B}_{2d,k} \mid \partial(\pi_1 p) = \partial(p) = k\}$. Here $\pi_1(\pi_2): R^d \times R^d = R^{2d} \to R^d$ is the projection on the first (second) factor. Finally put

$$\mathcal{B}_k = \{p \in \mathcal{B}^0_{2d,k} \mid \pi_2 p \subset \pi_1 p\}.$$

Then $\mathcal{B}_k \subset \mathcal{B}^0_{2d,k} \subset \mathcal{B}_{2d,k}$ and $\mathcal{B}^0_{2d,k}$ is an open subset of $\mathcal{B}_{2d,k}$. (All these sets are given the topology induced by $G_{2d,k}$.) Give $\bigcup_{k=0}^{d} \mathcal{B}_k$ the topology in which a set is open if and only if its intersection with each of the $\mathcal{B}_k$ is open.

Define a map $\mathfrak{s}: \mathcal{S} \to \bigcup_{k=0}^{d} \mathcal{G}_k$ by $\mathfrak{s}(S) = p$ where p is the essence of S.

LEMMA 4.3. (a) *$\mathfrak{s}: \mathcal{S} \to \bigcup_{k=0}^{d} \mathcal{G}_k$ is a homeomorphism.*

(b) *There is a continuous map $\mathfrak{g}: \bigcup_{k=0}^{d} \mathcal{G}_k \to \mathcal{G}$ which is one-one onto.*

(c) *There is a map $\rho_k: \mathcal{G}_{2d,k} \to \bigcup_{i \leq k} \mathcal{G}_i$ for each $k = 0, 1, \cdots, d$. $\rho_k \mid \mathcal{G}^0_{2d,k}: \mathcal{G}^0_{2d,k} \to \mathcal{G}_k$ is continuous (but ρ_k is not continuous on $\mathcal{G}_{2d,k} - \mathcal{G}^0_{2d,k}$).*

(d) *$\mathcal{G}$ is a submanifold of $G_{2d,d}$.*

Remarks. 1. From (a) and (b) follows the existence of a one-one correspondence $G: \mathcal{S} \to \mathcal{G}$ which is continuous (but not a homeomorphism). 2. As in Section 2 we say that a broken C^∞ curve X in R^d satisfies the boundary condition $S \in \mathcal{S}$ at s if X has no break at s and there is $x \in R^d$ such that $X(s) = S_1 x$, $X'(s) = S_2 x$. If $P = G(S) \in \mathcal{G}$, this is equivalent to demanding $(X(s), X'(s)) \in P$. It is in this sense the two definitions I and II are equivalent.

Proof of (a). $\mathfrak{s}$ is one-one, for if $\mathfrak{s}(S') = p = \mathfrak{s}(S)$, then $S'_1 = \pi_1 p = S_1$, and $(x, y) \in p$ implies $S'_2(x) = y = S_2(x)$. Thus $S'_2 \mid S'_1 = S'_2 \mid S_1 = S_2 \mid S_1$ which implies $S'_2 = S_2$ by (B3). $\mathfrak{s}$ is onto, for if $p \in \mathcal{G}_k$, let $\mathfrak{s}^{-1}(\mathrm{p}) = (S_1, S_2)$ where $S_1 = \pi_1 p$, and S_2 is defined on S_1 by $S_2 x = y$ where (x, y) is the unique element in p such that $\pi_1(x, y) = x$. S_2 is extended to all of R^d by the demand that (B3) hold. The defining property of $\mathcal{G}_k$ will guarantee that this pair (S_1, S_2) also satisfies (B2), and clearly $\mathfrak{s}(S_1, S_2) = p$.

To show the continuity of $\mathfrak{s}$ we first notice that any $S \in \mathcal{S}$ has a neighborhood U such that if $S' \in U$, then $\partial(S'_1) = \partial(S_1)$, and thus $\mathfrak{s}: U \to \mathcal{G}_k$ for some k. $\mathcal{S} \subset S(d) \times S(d)$ and hence the (matrix) norm $|\ |$ on $S(d)$ induces a metric on $\mathcal{S}$, compatible with the topology on $\mathcal{S}$.

Claim. For every sufficiently small $\epsilon > 0$ there is a ball $U_\delta \subset U$ of radius $\delta > 0$ about S, with the property that for every unit vector $u \in p = \mathfrak{s}(S)$ and for any $S' \in U_\delta$ there exists a vector $v' \in p' = \mathfrak{s}(S')$ such that $|u - v'| < \epsilon$. This claim implies the continuity of $\mathfrak{s}$ at S, as any frame spanning p can be approximated within ϵ by a frame spanning p'. To prove the claim, let $u = (x, S_2 x) \in p$, $|u| = 1$ and $S' \in U_\gamma$ where $U_\gamma \subset U$ is a ball of radius γ about S. Put $v' = (x', S'_2 x')$ when $x' = S'_1 x$. Then

$$\begin{aligned} | v' - u | &\leqq | x' - x | + | S'_2 x' - S_2 x | \\ &\leqq | S'_1 - S_1 | \, | x | + | S'_2 | \, | x' - x | + | S'_2 - S_2 | \, | x | \\ &\leqq \gamma + | S'_2 | \gamma + \gamma \leqq (2 + | S | + \gamma)\gamma \end{aligned}$$

which will be less than ϵ if $\gamma < \delta = \min (1, \epsilon/(3 + | S |))$.

We finally indicate why $\mathfrak{s}^{-1}$ is continuous. For any fixed $p \in \mathcal{G}_k$ and any sequence $p_n \to p$ in $\mathcal{G}_k$ we wish to show that $S_n = \mathfrak{s}^{-1}(p_n) \to S = \mathfrak{s}^{-1}(p)$ in $\mathcal{S}$. Since $p_n \to p$ clearly implies $\pi_1 p_n \to \pi_1 p$ in $G_{2d,k}$, $S_{1n} \to S_1$ in $S(d)$. Now $S_{2n} \to S_2$ in $S(d)$ if and only if $S_{2n} v \to S_2 v$ for all $v \in R^d$. But $v = w_n + w'_n = w + w'$ where $w_n = S_{1n} v$ and $w = S_1 v$, and hence

$$S_{2n} v = S_{2n} w_n + w'_n \to S_2 v = S_2 w + w'$$

if and only if $S_{2n} w_n \to S_2 w$. As $(w_n, S_{2n} w_n) \in p_n$, $(w, S_2 w) \in p$, and as $p_n \to p$ and $w_n \to w$, we conclude that $S_{2n} w_n \to S_2 w$, and hence that $S_{2n} \to S_2$.

Proof of (b). *If* $p \in \mathcal{G}_k$, p is extended to $P \in \mathcal{G}$ as follows. $P = \{ (x, y) + (0, z) \mid (x, y) \in p$ and $z \in (\pi_1 p)^{\perp} \}$. One checks directly that $P \in \mathcal{G}$, and this P is $\mathfrak{g}(p)$. $\mathfrak{g}$ is onto because every $P \in \mathcal{G}$ can be written uniquely in the form $P = \{ (x, y) + (0, z) \mid x, y \in \pi_1 P$ and $z \in (\pi_1 P)^{\perp} \}$. The same formula shows $\mathfrak{g}$ is one-one. The continuity of $\mathfrak{g}$ is now a consequence of the fact that if p' is sufficiently close to p in $\mathcal{G}_k$, then $p'_1 = \pi_1 p'$ is close to p_1 and $(p'_1)^{\perp}$ close to $(p_1)^{\perp}$ in $G_{d,k}$ and $G_{d,d-k}$ respectively.

Proof of (c). $\rho_k \colon \mathcal{G}_{2d,k} \to \bigcup_{i \leqq k} \mathcal{G}_i$ is defined by $\rho_k(q) = p$, where $p_1 = \pi_1 p = \pi_1 q$ and $p = \{ (x, y) \mid x \in p_1$ and $y = p_1 z$ when $(x, z) \in q\}$. (p is well defined, for if (x, z) and $(x, z') \in q$, then $(0, z - z') \in q$. Take any $x'' \in p_1$; then $(x'', z'') \in q$ for some $z'' \in R^d$. $q \in \mathcal{G}_{2d,k}$ implies

$$0 = J((x'', z''), (0, z - z')) = \langle x'', z - z' \rangle,$$

showing $y = p_1 z = p_1 z'$.) Similar arguments show $p \in \mathcal{G}_i$ for $i = \partial p_1 = \partial \pi_1 q$. The continuity of $\rho_k \mid \mathcal{G}^{0}_{2d,k}$ is shown by arguments of the type used to prove (a) and (b). ρ_k is not continuous at any point of $\mathcal{G}_{2d,k} - \mathcal{G}^{0}_{2d,k}$, and it is these discontinuities which are responsible for the "strong focal points."

Proof of (d). We refer to Edwards [7].

4.6. *Sprays and strong focal points.* For a fixed (α, s, S) the set $J^{*\alpha}{}_{Ss}$ is the linear space of all α-Jacobifields satisfying S (imposed) at s, which is of dimension d. $J^{\alpha}{}_{Ss}$ is the linear subspace of $J^{*\alpha}{}_{Ss}$ consisting of those Jacobifields X such that $X(s) \in S_1$ and $X'(s) = S_2 X(s)$, or equivalently

$(X(s), X'(s)) \in p$ where $p = \text{\textbackslash}(S)$ is the essence of S. Hence $\partial(J^\alpha{}_{Ss}) = \partial(S_1) = \partial(p)$.

Let $q \in \mathcal{G}_{2d,j}$. Because $\Phi^\alpha{}_{st} \in H$, $\Phi^\alpha{}_{st}(q) \in \mathcal{G}_{2d,j}$. For fixed (α, s, q) the curve $t \to \Phi^\alpha{}_{st}(q) : R^1 \to \mathcal{G}_{2d,j}$ is called the α-spray associated with q at s. It is a consequence of the continuity of the map Φ that the spray map Φ_j: $R^k \times R^1 \times R^1 \times \mathcal{G}_{2d,j} \to \mathcal{G}_{2d,j}$ is continuous. $\Phi_{\dot{\jmath}}$ is called the big spray. Next define Ψ: $R^k \times R^1 \times R^1 \times \mathcal{S} \to \mathcal{S}$ by $\Psi(\alpha, s, t, S) = \text{\textbackslash}^{-1}\rho_j\Phi_j(\alpha, s, t, \text{\textbackslash}(S))$ when $j = \partial(S_1)$. $\Psi(\alpha, s, t, S)$ is called the α-translate of the boundary condition S from s to t. It is this "pushing" which occurs in the index theorem of Ambrose, and it is completely determined for any t by the α-Jacobifields in $J^\alpha{}_{Ss}$. It is not transitive. We have broken it into a composition of several maps in order to prove the lemma below. The map ρ_j is not continuous. Thus Ψ will have certain discontinuities, the strong focal points. Explicitly, $(\alpha, s, t, S) \in R^k \times R^1 \times R^1 \times \mathcal{S}$ is a strong focal point if

$$\Phi_j(\alpha, s, t, \text{\textbackslash}(S)) \in \mathcal{G}_{2d,j} - \mathcal{G}^0{}_{2d,j}$$

when $\partial(S_1) = j$; or, equivalently, if $\partial(p) > \partial(E^\alpha{}_{st}(p))$ when $p = \text{\textbackslash}(S)$. The set of strong focal points in $R^k \times R^1 \times R^1 \times \mathcal{S}$ is closed, and for fixed (α, s, S) the numbers t for which (α, s, t, S) is a strong focal point are isolated, [2]. Lemma 4.3 (a) and (c) now imply

LEMMA 4.4. *If $(\alpha, s, t, S) \in R^k \times R^1 \times R^1 \times \mathcal{S}$ is not a strong focal point, then Ψ: $R^k \times R^1 \times R^1 \times \mathcal{S} \to \mathcal{S}$ is continuous in a neighborhood of (α, s, t, S).*

Remark. Because of the one-one correspondence G between $\mathcal{S}$ and $\mathcal{G}$, there is an analogous map

Ψ': $R^k \times R^1 \times R^1 \times \mathcal{G} \to \mathcal{G} : (\alpha, s, t, P) \to G(\Psi(\alpha, s, t, G^{-1}(P)))$. Strong focal points could be defined as above, but Ψ' is not continuous even away from these. The reason is that G^{-1}: $\mathcal{G} \to \mathcal{S}$ is not continuous.

The following example illustrates the discontinuities of Ψ'. If $e_1, \cdots, e_d$ is a frame in R^d, let $(\epsilon_1, \cdots, \epsilon_d, \delta_1, \cdots, \delta_d)$ be the frame of $R^d \times R^d$ where $\epsilon_i = (e_i, 0)$ and $\delta_i = (0, e_i)$. Let P be the d-plane spanned by $\delta_1, \cdots, \delta_d$, $P = L(\delta_1, \cdots, \delta_d)$. Then $\pi_1(P) = 0$, and $J \mid P \times P = 0$, so $P \in \mathcal{G}$. If $Q = L(\epsilon_1, \cdots, \epsilon_d)$, then $Q \in \mathcal{G}$ and $\pi_1(Q) = R^d$. For any real number r put $P_r = L(\delta_1 + r\epsilon_1, \cdots, \delta_d + r\epsilon_d)$. For all r near 0, P_r lies in a neighborhood of $P_0 = P$ in $\mathcal{G}$, and $\pi_1(P_r) = R^d$ when $r \neq 0$. Therefore the essence of P_r is all of P_r when $r \neq 0$, while the essence of P_0 is just $\{(0,0)\}$. Assume we deal with one curvature function only, which is the constant map R:

$R^1 \to S(d): t \to I_d$ where I_d is the identity map of R^d. The corresponding Jacobi equation is $X'' + X = 0$ (for a sphere) and the associated map

$$\Phi_{0t}: R^{2d} \to R^{2d}: (x, y) \to ((\cos t)x + (\sin t)y, (-\sin t)x + (\cos t)y).$$

In particular,

$$\Phi_{0\pi/2}: (x, y) \to (y, -x), \text{ and } \Phi_{0\pi/2}(P_r) = L(\epsilon_1 - r\delta_1, \cdots, \epsilon_d - r\delta_d).$$

Hence $\pi_1(\Phi_{0\pi/2}(P_r)) = R^d$, and for all r, including 0, the essence of $\Phi_{0\pi/2}(P_r)$ is $\Phi_{0\pi/2}(P_r)$. We conclude that the translate of P_r, $\Psi'(0, \pi/2, P_r)$, is $\Phi_{0\pi/2}(P_r)$ if $r \neq 0$, while $\Psi'(0, \pi/2, P_0) = P_0$. Since $\lim_{r\to 0} \Psi'(0, \pi/2, P_r) = Q$, we have shown that $\Psi'(0, \pi/2, \cdot)$ is not continuous at P_0. In fact, the formulas above show that $\Psi'(0, t, \cdot)$ is discontinuous at P_0 for all t in an interval about $\pi/2$.

4.7. *Continuity theorems.* Let

$$B = (\alpha, s, t, S, T) \in R^k \times R^1 \times R^1 \times \mathcal{S} \times \mathcal{S}.$$

Call r a *conjugate point of* B if r is a conjugate point of the pair (S, T) in the Ambrose sense [2], where S and T are imposed at s and t respectively, and the index form I_B is defined on $\Xi(B) \times \Xi(B)$ by

$$I_B(X, Y) = f(\alpha)\Big[\int_s^t (\langle X', Y'\rangle - \langle R^\alpha X, Y\rangle) \\ - \langle T_2 X(t), Y(t)\rangle + \langle S_2 X(s), Y(s)\rangle\Big].$$

$f > 0$ is continuous, and $\Xi(B) = \{X \in \Xi(s, t) \mid X(s) \in S_1 \text{ and } X(t) \in T_1\}$. $\Xi(B)$ does not depend on α. $n(r, B)$ is the order of r as a conjugate point of B, and $i(B)$ the index of I_B.

THEOREM 4.1. *Select any* $B = (\alpha, s, t, S, T) \in R^k \times R^1 \times R^1 \times \mathcal{S} \times \mathcal{S}$ *and suppose* $s < r < t$. *Suppose further that there are no conjugate points of* B *in the interval* $[t_1, t_2]$ *except (possibly)* r, *where* $s < t_1 < r < t_2 < t$, *and that* $J^{*\alpha}{}_{Ss} \cap J^{\alpha}{}_{Tt} = \{0\}$. *Then there exists a neighborhood* U *of* B *such that for any* $\bar{B} \in U$

$$\sum_{t_1 < v < t_2} n(v, \bar{B}) = n(r, B).$$

THEOREM 4.2. *If* $B = (\alpha, s, t, S, T)$ *is such that* $n(t, B) = 0$, *then there exists a neighborhood* U *of* B *such that for any* $\bar{B} \in U$, $i(\bar{B}) = i(B)$ *and* $n(t, \bar{B}) = n(t, B) = 0$.

Proof of Theorem 4.1, *assuming Theorem* 4.2. Given $B = (\alpha, s, t, S, T)$ the map: $(\alpha, t, T) \to \Psi(\alpha, t, t_i, T) = T_i$: $R^k \times R^1 \times \mathcal{S} \to \mathcal{S}$ is continuous in a

neighborhood V_i of (α, t, T) $(i=1,2)$ by Lemma 4.4, because the assumption of Theorem 4.1 implies in particular that (α, t, t_i, T) is not a strong focal point. Hence the corresponding map:

$$R^k \times R^1 \times R^1 \times \mathcal{S} \times \mathcal{S} \to R^k \times R^1 \times R^1 \times \mathcal{S} \times \mathcal{S}:$$
$$(\alpha, s, t, S, T) \to B_i = (\alpha, s, t_i, S, T_i)$$

is continuous in a neighborhood U_i of B. Combining this with the statement of Theorem 4.2, we obtain the existence of a neighborhood U such that for any $\bar{B} \in U$, $i(\bar{B}_i) = i(B_i)$ and $n(\bar{B}_i) = n(B_i)$. Then from the Ambrose index theorem [2] follows

$$n(r, B) = i(B_2) - i(B_1) = i(\bar{B}_2) - i(\bar{B}_1) = \sum_{t_1 < v < t_2} n(v, \bar{B})$$

for any $\bar{B} \in U$.

Proof of Theorem 4.2. $B = (\alpha, s, t, S, T)$ is fixed with $n(t, B) = 0$. Suppose the dimensions of S_1 and T_1 are m and n respectively. By Lemma 4.2 there is a neighborhood N_1 of (α, s, t) in $R^k \times R^1 \times R^1$ and a sequence $\{t_i\}$, $i = 1, \cdots, p$, $p > 2$, which is $\bar{\alpha}$-normal for the interval $[\bar{s}, \bar{t}]$ for every $(\bar{\alpha}, \bar{s}, \bar{t}) \in N_1$. Let $\mathcal{B}(B)$ be the subspace of $\Xi(B)$ consisting of broken α-Jacobifields relative to the α-normal sequence $\{t_i\}$. Then the index and nullity of I_B on $\Xi(B) \times \Xi(B)$ equal the index and nullity of $I_B \mid \mathcal{B}(B) \times \mathcal{B}(B)$. We denote this restriction also by I_B, and prove (a) there exist a neighborhood V of B and for evey $\bar{B} \in V$ an isomorphism $\Theta_{\bar{B}}\colon S_1 \times R^{pd} \times T_1 \to \mathcal{B}(\bar{B})$; (b) the bilinear symmetric form $\Theta_{\bar{B}}^*(I_{\bar{B}})$ depends continuously on $\bar{B} \in V$.

(a) and (b) imply Theorem 4.2 because the eigenvalues of $\Theta_{\bar{B}}^*(I_{\bar{B}}) = J_{\bar{B}}$ will depend continuously on $\bar{B}$. Thus the assumption that $n(t, B) = 0$ implies that the eigenvalues of J_B are strictly positive or negative and hence will stay so in some neighborhood U of B. Hence the index $i(J_B) = i(J_{\bar{B}})$ and the nullity $n(J_B) = 0 = n(J_{\bar{B}})$ for all $\bar{B} \in U$, and this yields Theorem 4.2 via (a).

Proof of (a). Choose a neighborhood V of B such that $\bar{B} \in V$ implies that $\{t_i\}$ is an $\bar{\alpha}$-normal sequence for $[\bar{s}, \bar{t}]$ and that $\partial(\bar{S}_1) = \partial(S_1)$ and $\partial(\bar{T}_1) = \partial(T_1)$. If V is sufficiently small, $\bar{S}_1$ and $\bar{T}_1$ are close to S_1 and T_1 respectively in $G_{d,m}$ and $G_{d,n}$, and there are orthogonal projections $p_{\bar{S}S}\colon \bar{S}_1 \to S_1$ and $p_{\bar{T}T}\colon \bar{T}_1 \to T$ which are isomorphisms. If

$$x = (x_0, x_1, \cdots, x_p, x_{p+1}) \in S_1 \times (R^d)^p \times T_1,$$

define $\Theta_{\bar{B}}\colon S_1 \times (R^d)^p \times T_1 \to \mathcal{B}(\bar{B})$ by $\Theta_{\bar{B}}(x) =$ unique broken $\bar{\alpha}$-Jacobifield X such that X has breaks only at the t_i and $X(\bar{s}) = p_{\bar{S}S}^{-1}(x_0), X(t_i) = x_i$ $(i = 1, \cdots, p)$ and $X(\bar{t}) = p_{\bar{T}T}^{-1}(x_{p+1})$. This map is an isomorphism.

Proof of (b). Choose orthogonal bases $a_1, \cdots, a_m, b_{11}, \cdots, b_{pd}, c_1, \cdots, c_n$ of S_1, $(R^d)^p$ and T_1 respectively, and denote by the same letters the induced basis in $S_1 \times (R^d)^p \times T_1$. The form $J_{\bar{B}}$ is continuous if its coefficients relative to the basis $a_1, \cdots, c_n$ of $S_1 \times (R^d)^p \times T_1$ are continuous. Let $\bar{A}_1 = \Theta_{\bar{B}}(a_1), \cdots$. We must show the functions $J_{\bar{B}}(a_i, a_j) = I_{\bar{B}}(\bar{A}_i, \bar{A}_j)$, $J_{\bar{B}}(a_i, b_{jk}) = I_{\bar{B}}(\bar{A}_i, \bar{B}_{jk}), \cdots$ are continuous in $\bar{B}$. For this observe that if $X, Y \in \mathfrak{B}(\bar{B})$, then

$$I_{\bar{B}}(X,Y) = f(\bar{\alpha})\Big(\sum_{i=1}^{p}\langle X'(t_i^-) - X'(t_i^+), Y(t_i)\rangle + \langle X'(\bar{t}) - \bar{T}_2 X(\bar{t}), Y(\bar{t})\rangle - \langle X'(\bar{s}) - \bar{S}_2 X(\bar{s}), Y(\bar{s})\rangle\Big).$$

Hence $J_{\bar{B}}(a_i, a_j) = f(\bar{\alpha})\langle \bar{S}_2 \bar{A}_i(\bar{s}) - \bar{A}_i'(\bar{s}), \bar{A}_j(\bar{s})\rangle$ is continuous if $\bar{A}_i(\bar{s})$ and $\bar{A}_i'(\bar{s})$ are continuous in $\bar{B}$. But $\bar{A}_i$ (restricted to $[\bar{s}, t_1]$) is the unique $\bar{\alpha}$-Jacobifield which at $\bar{s}$ is $p_{\bar{S}S}^{-1}(a_i)$ and at t_1 is 0. Hence $\bar{A}_i(\bar{s}) = p_{\bar{S}S}^{-1}(a_i)$ and $\bar{A}_i'(\bar{s}) = F'^{\bar{\alpha}}_{t_1\bar{s}} \circ (F^{\bar{\alpha}}_{t_1\bar{s}})^{-1} \circ p_{\bar{S}S}^{-1}(a_i)$ which do depend continuously on $\bar{B}$ in V. Similar arguments will show the continuity of $J_{\bar{B}}(a_i, b_{jk}), \cdots$.

Remarks. 1. If (s,t) is not an α-conjugate pair the proof of Theorem 4.2 could be simplified by the use of Theorem 3.3.

2. The topology on the boundary conditions used in the above theorems was the (geometric) one of the space $\mathcal{S}$. The following example shows there is no analog of Theorem 4.2 if we use the (coarser) topology of $\mathfrak{S}$. For each positive real number r consider the self adjoint boundary conditions $S_r = (I_d, (-1/r)I_d)$ where $I_d \colon R^d \to R^d$ is the identity map, and $T = (0, I_d)$, imposed at 0 and 1 respectively. Let $R \colon R^1 \to S(d)$ be identically 0. The corresponding index form I_r on $\Xi(R^d, 0) \times \Xi(R^d, 0)$ is

$$I_r(X,Y) = \int_0^1 \langle -X'', Y\rangle + \sum \langle X'(t_i^-) - X'(t_i^+), Y(t_i)\rangle - \langle (1/r)X(0) + X'(0), Y(0)\rangle.$$

(This boundary problem arises from a straight line segment in R^{d+1} which at its left point is perpendicular to a sphere of radius r with opening to the right, and on the right ends in a point.) The index $i(I_r) = d$ and the nullity $n(I_r) = 0$, if $0 < r < 1$, as is seen by finding the focal points of S_r. As $r \to 0$ the plane $P_r = G(S_r)$ which is the boundary condition S_r considered as an element of $\mathfrak{S}$, approaches the limiting position $\{0\} \times R^d = Q = G(T)$. Hence if $B_r = (0, 1, P_r, Q)$ and $B = (0, 1, Q, Q)$ then $B_r \to B$ in $R^1 \times R^1 \times \mathfrak{S} \times \mathfrak{S}$. But $i(I_B) = 0 = n(I_B)$, and thus $i(I_B) \neq \lim_{r \to 0} i(I_r) = d$, even though $n(I_B) = 0$.

4.8. *Ordinary conjugate points and continuity.* In the simplest boundary problem, the fixed ends case, the continuity theorem 4.1 can be strengthened. If m and n are conjugate points along a given geodesic σ, and if m is moved sufficiently slightly to the right along σ to a position m', then m' has a conjugate point n' which lies to the right of n on σ. This fact was essentially proved by Morse [11]. The proof is based on the index theorem for this case and the continuity of the index form, and we do not repeat it here. We only give a precise statement.

THEOREM 4.3. (Morse). *Suppose (s,t) is a conjugate pair of order k along the geodesic σ. For any given $\epsilon > 0$ such that there are no conjugate points of s in $(t-\epsilon, t+\epsilon)$ except t, there is a $\delta > 0$ such that if $x \in (s, s+\delta)$, there are precisely k conjugate points of x in $(t, t+\epsilon)$.*

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE,
UNIVERSITY OF CALIFORNIA, BERKELEY.

REFERENCES.

[1] W. Ambrose, "The Cartan structural equations in classical Riemannian geometry," *Indian Mathematical Society*, vol. 24 (1960), pp. 23-76.

[2] ———, "The index theorem in Riemannian geometry," *Annals of Mathematics*, vol. 73 (1961), pp. 49-86.

[3] E. Artin, *Geometric Algebra*, Interscience Tracts in Pure and Applied Mathematics, 1957.

[4] R. Bott, "On the iteration of closed geodesics and the Sturm intersection theory," *Communications on Pure and Applied Mathematics*, vol. 9 (1956), pp. 171-206.

[5] ———, "The stable homotopy of the classical groups," *Annals of Mathematics*, vol. 70 (1959), pp. 313-317.

[6] C. Chevalley, *Theory of Lie Groups*, Princeton University Press (1946).

[7] H. M. Edwards, Jr., *Applications of Intersection Theory to Boundary Value Problems*, Thesis, Harvard University (1961).

[8] N. Levinson and E. A. Coddington, *Theory of Ordinary Differential Equations*, McGraw-Hill (1955).

[9] M. Morse, *The Calculus of Variations in the Large*, Colloquium Publications, American Mathematical Society (1934).

[10] ———, "Closed extremals," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 15 (1929), pp. 856-859.

[11] ———, "A generalization of the Sturm separation and comparison theorems in n-space," *Mathematische Annalen*, vol. 103 (1930), pp. 52-69.

[12] ———, "The index theorem in the calculus of variations," *Duke Mathematical Journal*, vol. 4 (1938), pp. 231-246.

[13] ———, "Generalized concavity theorems," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 21 (1935), pp. 359-362.

[14] I. M. Singer, *Lecture Notes* by E. M. Brown, MIT, Cambridge (1962).

MAXIMALE HOLOMORPHE UND MEROMORPHE ABBILDUNGEN, I.*

Von Karl Stein.†

Einleitung. Es seien X, Y, Y_0 komplexe Räume, $f: X \to Y$ und $f_0: X \to Y_0$ seien holomorphe Abbildungen. f_0 heisst von f *strikt abhängig*, wenn die durch die Zuordnung $x \to (f(x), f_0(x))$, $x \in X$, bestimmte holomorphe Abbildung $(f, f_0): X \to Y \times Y_0$ in jedem Punkte von X den gleichen lokalen Rang wie f besitzt; f und f_0 heissen *im engeren Sinne verwandt*, wenn f_0 von f und f von f_0 strikt abhängig ist. Die Abbildung f heisst *maximal*, wenn sie surjektiv ist und wenn jede von f strikt abhängige holomorphe Abbildung $f': X \to Y'$ eine Faktorisierung $f' = \alpha(f') \circ f$ mit einer holomorphen Abbildung $\alpha(f'): Y' \to Y$ gestattet. Sind f und f_0 im engeren Sinne verwandt und ist f maximal, so wird das Paar (f, Y) als eine *komplexe Basis* zur Abbildung f_0 bezeichnet.

Es hat sich gezeigt, dass die Begriffe *maximale holomorphe Abbildung* und *komplexe Basis zu einer holomorphen Abbildung* in mancher Hinsicht mit Nutzen verwendet werden können (vgl. [2, 13, 21, 23, 28]). Die wichtige Frage nach der Existenz einer komplexen Basis zu einer vorgegebenen holomorphen Abbildung ist in verschiedenen Arbeiten behandelt worden [15, 28, 29, 30].

Die vorliegende Arbeit enthält weitere Beiträge zur Theorie der maximalen Abbildungen. Es wird eine weitere hinreichende Bedingung für die Existenz komplexer Basen zu gegebenen holomorphen Abbildungen angegeben, und es werden Eigenschaften maximaler holomorpher Abbildungen sowie spezielle Fälle diskutiert. Die Betrachtungen werden ferner auf meromorphe Abbildungen ausgedehnt. Es werden einige allgemeine Sätze über maximale meromorphe Abbildungen und im Zusammenhang damit Aussagen über Systeme abhängiger meromorpher Abbildungen bewiesen. Im Rahmen dieser Betrachtungen lässt sich durch Spezialisierung u. a. ein Satz von W. Thimm [34, 36] über die algebraische Abhängigkeit von abhängigen meromorphen Funktionen mit gemeinsamen Unbestimmtheitsstellen gewinnen.[1]

* Received January 4, 1963.

† The author was supported by AFOSR through the European Office of Aerospace Research, US Air Force Grant No. AF-EOAR-61-50.

[1] Dieser Satz wurde auch von R. Remmert in [24], als Sonderfall einer allgemeineren Aussage über abhängige meromorphe Funktionen, neu gewonnen. Von P. Roquette wurde kürzlich der Satz von Thimm im Rahmen der lokalen Algebra formuliert und bewiesen.

Im einzelnen gliedert sich der vorliegende Teil I der Arbeit wie folgt:

In § 1 wird zunächst an wichtige Begriffe und Sachverhalte aus der Theorie der komplexen Räume erinnert (der Begriff *komplexer Raum* ist hier stets im Sinne der von J. P. Serre [27] gegebenen Definition gemeint). Nach einigen vorbereitenden Aussagen wird bewiesen (Satz I), dass zu einer holomorphen Abbildung $f: X \to Y$ eines irreduziblen komplexen Raumes X immer dann eine komplexe Basis existiert, wenn es in X eine analytische Menge A gibt, derart, dass die Beschränkung $f \mid A$ eigentlich ist und den gleichen globalen Rang wie f besitzt. Spezialisierung auf den Fall $A = X$ ergibt, dass zu einer eigentlichen holomorphen Abbildung f stets eine mit f im engeren Sinne verwandte maximale holomorphe Abbildung existiert. Diese letzte Aussage, und zwar in verschärfter Form, folgt unmittelbar auch aus einem Resultat von H. Cartan in [8]. Auf weitere Literaturzusammenhänge und auf Ausdehnungsmöglichkeiten der Aussagen wird am Schluss von § 1 hingewiesen.

§ 2 handelt von Eigenschaften maximaler holomorpher Abbildungen. Es wird gezeigt, dass eine maximale holomorphe Abbildung stets einen Epimorphismus der Fundamentalgruppen der beteiligten komplexen Räume induziert. Existiert zu einer vorgegebenen holomorphen Abbildung $f: X \to Y$ eine komplexe Basis (f^*, Y^*), so ist für manche Anwendungen die Frage von Bedeutung, wann der Raum Y^* kompakt ist. Es werden einige Fälle betrachtet, in denen dies eintritt. Weiter wird eine Konsequenz aus Satz I angegeben (Proposition 2.3.1), die Ringe holomorpher Funktionen auf X und Y betrifft. Schliesslich wird auf Eigenschaften besonderer maximaler Abbildungen von Holomorphiehüllen komplexer Räume eingegangen.

Maximale *meromorphe* Abbildungen werden im nachfolgenden Teil II dieser Arbeit behandelt. Zum Beweis der Hauptresultate wird eine Reihe von Vorbereitungen, unter anderem über holomorphe Korrespondenzen, erforderlich sein.

1. Maximale holomorphe Abbildungen. Eine Existenzaussage.

1.1. Wir beginnen mit Bemerkungen zur Terminologie und erinnern an wichtige Aussagen. Hinsichtlich der nicht näher erläuterten Begriffe aus der Theorie der komplexen Räume und ihrer wesentlichen Eigenschaften sei insbesondere auf [3, 6, 7, 12, 26] verwiesen.

Unter einem komplexen Raum X verstehen wir in dieser Arbeit stets einen komplexen Raum im Sinne von J. P. Serre [27] [2]; alle vorkommenden

[2] Eine allgemeinere Definition des Begriffes *komplexer Raum* wurde von H. Grauert

komplexen Räume werden als nichtleer und endlich-dimensional vorausgesetzt. Die komplexe Dimension von X im Punkte $x \in X$ wird mit $\dim_x X$ bezeichnet, die komplexe Dimension von X schlechthin ist $\dim X := \max_{x \in X} \dim_x X$. Ist A eine nichtleere lokal-analytische Menge in X, so induziert die komplexe Struktur von X eine komplexe Struktur auf A; A kann und soll immer auch als ein mit dieser Struktur versehener komplexer Raum aufgefasst werden und heisst dann ein komplexer Unterraum von X. Ist $f: X \to Y$ eine holomorphe Abbildung, so ist die Beschränkung $f \mid A: A \to Y$ ebenfalls holomorph. Die komplexe Codimension von A im Punkte $x \in A$ ist

$$\operatorname{codim}_x A := \dim_x X - \dim_x A,$$

die komplexe Codimension von A schlechthin ist erklärt als die Zahl $\operatorname{codim} A := \min_{x \in A} \operatorname{codim}_x A$.

Unter dem (*lokalen*) *Rang* $r_x(f)$ der holomorphen Abbildung $f: X \to Y$ im Punkte $x \in X$ wird die Codimension der Faser $f^{-1}(f(x))$ im Punkte x verstanden; die Zahl $r(f) := \max_{x \in X} r_x(f)$ heisst der (*globale*) *Rang* von f in X. Ist $r_x(f) = \dim_x X$ für alle $x \in X$ (d.h. sind alle Fasern $f^{-1}(f(x))$ diskret), so heisst f *nirgends entartet*. Sind $f_\lambda: X \to Y_\lambda$ $(\lambda = 1, \cdots, l)$ weitere holomorphe Abbildungen von X, so nennen wir die durch die Zuordnung $x \to (f_1(x), \cdots, f_l(x))$ bestimmte holomorphe Abbildung

$$\Phi: X \to Y_1 \times \cdots \times Y_l$$

die *Verbindung* von $f_1, \cdots, f_l$ und schreiben $\Phi = (f_1, \cdots, f_l)$. Die holomorphe Abbildung $f_0: X \to Y_0$ heisst von der Abbildung f *strikt abhängig*, wenn $r_x(f) = r_x((f, f_0))$ für jeden Punkt $x \in X$ gilt; f und f_0 heissen *im engeren Sinne verwandt*, wenn f von f_0 und f_0 von f strikt abhängig ist. Wir nennen ferner f_0 schlechthin von f *abhängig*—oder auch von f *abhängig im weiteren Sinne*—wenn für jede irreduzible Komponente X' von X $r(f \mid X') = r((f \mid X', f_0 \mid X'))$ gilt; f und f_0 heissen *verwandt im weiteren Sinne* oder auch *analytisch verwandt*, wenn f von f_0 und f_0 von f im weiteren Sinne abhängig ist.[3] Es gilt: f_0 ist genau dann von f strikt abhängig, wenn f_0 auf den *Niveaumengen* von f (d.h. auf den zusammenhängenden Komponenten der Fasern $f^{-1}(f(x))$, $x \in X$, von f) konstant ist; f_0 ist genau dann von f im weiteren Sinne abhängig, wenn für jede irreduzible Komponente X'

in [10] gegeben; komplexe Räume im Sinne der Definition von J. P. Serre werden dort als reduziert bezeichnet.

[3] Wir benutzen hier eine gegenüber [26], [29] und [30] etwas veränderte Terminologie.

von X $f_0 \mid X'$ auf den Niveaumengen von $f \mid X'$ der kleinsten vorkommenden Dimension konstant bleibt. Ist $r(f_0) = 1$ und f_0 von f abhängig im weiteren Sinne, so ist f_0 von f sogar strikt abhängig.

Wir sagen, die holomorphe Abbildung $f: X \to Y$ *majorisiere* die holomorphe Abbildung $f_0: X \to Y_0$, wenn eine holomorphe Abbildung $\alpha: Y \to Y_0$ mit $f_0 = \alpha \circ f$ existiert. Die Abbildung f heisse *maximal*, wenn sie surjektiv ist und wenn sie jede von f strikt abhängige holomorphe Abbildung majorisiert. Sind f_0 und f im engeren Sinne verwandt und ist f maximal, so nennen wir das Paar (f, Y) eine *komplexe Basis zur Abbildung* f_0. Jede weitere komplexe Basis (f', Y') zur Abbildung f_0 ist mit (f, Y) *biholomorph äquivalent*, d. h. es gibt eine biholomorphe Abbildung $\phi: Y \to Y'$, sodass $f' = \phi \circ f$ gilt.

Ist die holomorphe Abbildung $f: X \to Y$ eigentlich, nirgends entartet, surjektiv und wird jede irreduzible Komponente von X vermöge f auf eine irreduzible Komponente von Y abgebildet, so heisst f eine *eigentliche Überlagerungsabbildung* und das Tripel (X, f, Y) eine *analytische Überlagerung von Y durch X*. Jede Faser $f^{-1}(f(x))$ ist dann endlich; ferner gibt es in X bezw. in Y je eine nirgendsdichte analytische Menge M_X bezw. M_Y, derart, dass $f(M_X) = M_Y$ und $f^{-1}(M_Y) = M_X$ gilt und dass die durch Beschränkung von f bestimmte eigentliche surjektive holomorphe Abbildung $f': X - M_X \to Y - M_Y$ lokal biholomorph ist. Ist $Y = \bigcup_i Y^{(i)}$ die Zerlegung von Y in irreduzible Komponenten, so besteht für $y \in Y^{(i)} - M_Y$ die Menge $f^{-1}(y)$ jeweils aus der gleichen (endlichen) Anzahl $b^{(i)}$ von Punkten; $b^{(i)}$ heisst die *Blätterzahl von* (X, f, Y) *über* $Y^{(i)}$ und $b := \sup_i b^{(i)}$ die *Blätterzahl von* (X, f, Y) schlechthin (b kann endlich oder unendlich sein).

Der komplexe Raum X heisst *im Punkte $x \in X$ normal*, wenn der Halm über x der Strukturgarbe der holomorphen Funktionskeime ein ganz abgeschlossener Integritätsring ist; X heisst *normal* schechthin, wenn X in jedem seiner Punkte normal ist. Zu jedem komplexen Raum X existiert eine analytische Überlagerung $('X, \nu, X)$ mit der Blätterzahl 1 durch einen normalen komplexen Raum $'X$; eine solche analytische Überlagerung heisst eine *Normalisierung* von X (vgl. hierzu [7, 12, 18, 19]). Zwei Normalisierungen $('X, \nu, X)$ und $('X_1, \nu_1, X)$ sind stets biholomorph äquivalent, d. h. es gibt eine biholomorphe Abbildung $\mu: 'X \to 'X_1$ mit $\nu = \nu_1 \circ \mu$. Eine Normalisierung $('X, \nu, X)$ ist speziell auch eine eigentliche Modifikation von X (vgl. hierzu § 4 dieser Arbeit).

1.2. Wir stellen einige im folgenden benötigte einfache Aussagen zusammen.

PROPOSITION 1.2.1. *Sei $f: X \to Y$ eine holomorphe Abbildung, A eine nichtleere lokal-analytische Menge in X. Dann gilt $r(f \mid A) \leqq r(f)$.*

Beweis. Es genügt, X und A als irreduzibel anzunehmen. Sei x ein auf A liegender Punkt von X und U eine offene Umgebung von x in X mit abzählbarer Topologie; dann ist $U' := U \cap A$ eine in A offene Menge mit abzählbarer Topologie. Die Bildmengen $f(U)$ und $(f \mid A)(U') = f(U')$, aufgefasst als topologische Unterräume von Y, haben jeweils die topologische Dimension $2r(f)$ bezw. $2r(f \mid A)$ (vgl. [26], Satz 1). Da $f(U')$ in $f(U)$ enthalten ist, muss notwendig $r(f \mid A) \leqq r(f)$ gelten.

PROPOSITION 1.2.2. *Sei X ein irreduzibler komplexer Raum, A eine nichtleere analytische Menge in $X, f: X \to Y$ eine holomorphe Abbildung, derart, dass $f \mid A: A \to Y$ eigentlich und $r(f) = r(f \mid A)$ ist. Dann ist $f(A) = f(X)$, und $f(X)$ ist eine $r(f)$-dimensionale irreduzible analytische Menge in Y.*[4]

Beweis. X darf als normal vorausgesetzt werden (nötigenfalls ist zu einer Normalisierung $('X, \nu, X)$ überzugehen und $'X, 'A := \nu^{-1}(A)$, $'f := f \circ \nu$: $'X \to Y$ an Stelle von X, A, f zu betrachten). Weiter darf angenommen werden, dass A irreduzibel ist.

Sei X' bezw. A' die Menge derjenigen Punkte von X bezw. A, in denen der lokale Rang von f bezw. von $f \mid A$ gleich dem globalen Rang $r(f)$ bezw. $r(f \mid A)$ ist; X' und A' sind jeweils in X bezw. in A nichtleer, offen und zusammenhängend. Wir behaupten zunächst, dass der Durchschnitt $A'' := X' \cap A'$ nichtleer ist. Träfe dies nicht zu, so würde $X - X' \supset A$ gelten. $X - X' =: N$ ist eine analytische Menge in X (vgl. [22]); sei N_0 eine von denjenigen irreduziblen Komponenten von N, die A enthalten. Nun gilt $r(f) \geqq r(f \mid N_0) + 2$; andererseits folgt aus Proposition 1.2.1, das

$$r(f \mid N_0) \geqq r((f \mid N_0) \mid A) = r(f \mid A)$$

ist. Also wäre $r(f) > r(f \mid A)$, im Widerspruch zur Voraussetzung.

Sei jetzt A^* die Menge derjenigen Punkte von A'', in denen A irreduzibel ist; A^* ist eine in A offene und dort dichte Menge. Wird der (normale) komplexe Raum der analytische Primkeime von Y mit $\mathfrak{S}(Y)$ bezeichnet (vgl. [6, 7, 30]) und ist $\tau_Y: \mathfrak{S}(Y) \to Y$ die Projektion von $\mathfrak{S}(Y)$ auf Y, so gibt es holomorphe Abbildungen $f': X' \to \mathfrak{S}(Y)$ und $f^*: A^* \to \mathfrak{S}(Y)$ mit $f \mid X' = \tau_Y \circ f'$ bezw. $f \mid A^* = \tau_Y \circ f^*$, und zwar sind jeweils $f'(X')$ und

[4] Vgl. hierzu [24], p. 894 (Lemma), wo für Räume mit abzählbarer Topologie eine weitergehende Aussage bewiesen ist.

$f^*(A^*)$ nichtleere offene zusammenhängende Mengen in $r(f)$-dimensionalen Komponenten G' bezw. G^* von $\mathfrak{G}(Y)$. Da $f \mid A$ als eigentlich und $r(f) = r(f \mid A)$ vorausgesetzt ist, ist nach R. Remmert [22] die Menge $B := f(A) \subset Y$ eine $r(f)$-dimensionale irreduzible analytische Menge in Y; $(G^*, \tau_Y \mid G^*, B)$ ist eine Normalisierung des komplexen Unterraumes B von Y. Sei nun x_0 irgendein Punkt von A^*, dann ist der durch $f^*(x_0)$ repräsentierte $r(f)$-dimensionale analytische Primkeim notwendig in dem durch $f'(x_0)$ repräsentierten $r(f)$-dimensionalen analytischen Primkeim enthalten. Also ist $f^*(x_0) = f'(x_0)$ *und* infolgedessen $G' = G^*$. Dies bedeutet, dass $f(X')$ in $f(A) = B$ enthalten ist; da aber X' in X dicht liegt und B als analytische Menge abgeschlossen in Y ist, muss dann auch $f(X) \subset f(A)$ gelten. Andererseits ist wegen $A \subset X$ natürlich $f(A) \subset f(X)$. Mithin gilt $f(A) = f(X)$.

PROPOSITION 1.2.3. *X, Y seien nichtleere topologische Räume; Y sei lokal-kompakt; $f\colon X \to Y$ sei eine Abbildung, derart, dass jede Faser $f^{-1}(f(x))$, $x \in X$, in jeweils endlich viele zusammen hängende Komponenten zerfällt. Für jede positive ganze Zahl l sei M_l die Menge derjenigen Punkte $y \in Y$, für welche $f^{-1}(y)$ höchstens l zusammenhängende Komponenten besitzt. Dann gibt es eine ganze Zahl l_0, sodass M_l für $l \geqq l_0$ in einer offenen Menge $E \neq \emptyset$ von Y dicht liegt.*

Da $M_l \subset M_{l+1}$ und $\bigcup_l M_l = Y$ gilt, folgt die Aussage unmittelbar aus dem Theorem von Baire.

1.3. Zum Beweise des folgenden Satzes I benötigen wir

PROPOSITION 1.3.1. *Voraussetzung. Sei $\mathfrak{J}$ eine vermöge einer Ordnungsrelation $\leqq$ gerichtete Menge; Y_i $(i \in \mathfrak{J})$ seien komplexe Räume. Zu jedem Paar $i, j \in \mathfrak{J}$ mit $j \leqq i$ gebe es eine eigentliche Überlagerungsabbildung $\alpha_{ij}\colon Y_i \to Y_j$, derart, dass für $k \in \mathfrak{J}$ mit $k \leqq j$ stets $\alpha_{ik} = \alpha_{jk} \circ \alpha_{ij}$ gilt. Es existiere ferner ein komplexer Raum $\check{Y}$ und zu jedem $i \in \mathfrak{J}$ eine eigentliche Überlagerungsabbildung $\phi_i\colon \check{Y} \to Y_i$, sodass für $j \leqq i$ stets $\phi_j = \alpha_{ij} \circ \phi_i$ ist. Weiter sei $\check{U}$ eine nichtleere relativ kompakte offene Teilmenge von $\check{Y}$; es sei $U_i := \phi_i(\check{U})$ gesetzt (U_i ist relativ kompakt und offen in Y_i), und es bezeichne $'\alpha_{ij}\colon U_i \to U_j$ die durch Beschränkung von α_{ij} bestimmte Abbildung.*

Behauptung. Es gibt einen Index $i_0 \in \mathfrak{J}$, sodass $'\alpha_{ij}$ für $i_0 \leqq j \leqq i$ biholomorph ist.

Der Beweis wird mittels vollständiger Induktion nach der Dimension des Raumes $\check{Y}$ geführt. Die Behauptung ist offensichtlich richtig, falls $\dim \check{Y} = 0$. Angenommen, sie sei bewiesen für Dimensionen bis $n-1$

$(n>0)$. Sei jetzt $\dim \check{Y}=n$. Wir wählen eine relativ kompakte offene Teilmenge $\check{U}^*$ von $\check{Y}$, in der $\check{U}$ relativ kompakt enthalten ist; es werde $U_i^*:=\phi_i(\check{U}^*)$ gesetzt, und es sei $\alpha_{ij}^*: U_i^* \to U_j^*$ (für $j \leqq i$) jeweils die durch Beschränkung von α_{ij} bestimmte Abbildung. Wir weisen zunächst die Existenz eines Index $j_0 \in \mathfrak{J}$ nach, derart, dass für $j_0 \leqq j \leqq i$ die α_{ij}^* sämtlich Homöomorphismen sind. $\check{U}^*$ hat nur mit endlich vielen irreduziblen Komponenten von $\check{Y}$, etwa mit $\check{Y}^{(\mu)}$ $(\mu=1,\cdots,m)$ gemeinsame Punkte. Sei $Y_i^{(\mu)}:=\phi_i(\check{Y}^{(\mu)})$ und $\phi_i^{(\mu)}: \check{Y}^{(\mu)} \to Y_i^{(\mu)}$ die durch Beschränkung von ϕ_i bestimmte Abbildung; dann ist jeweils $(Y^{(\mu)}, \phi_i^{(\mu)}, Y_i^{(\mu)})$ eine analytische Überlagerung, etwa der Blätterzahl $b_i^{(\mu)}$. Da für $j \leqq i$ sicher $b_j^{(\mu)} \leqq b_i^{(\mu)}$ gilt, gibt es einen Index j_1, sodass für $j_1 \leqq j \leqq i$ und $\mu=1,\cdots,m$ stets $b_j^{(\mu)}=b_i^{(\mu)}$ ist. Es folgt, dass in $\check{Y}^{(\mu)}$ und $Y_i^{(\mu)}$ $(j_1 \leqq i)$ nirgends dichte analytische Mengen $\check{M}^{(\mu)}$ bezw. $M_i^{(\mu)}$ existieren, sodass

$$\phi_i(\check{M}^{(\mu)})=M_i^{(\mu)}, \qquad \phi_i(\check{Y}^{(\mu)}-\check{M}^{(\mu)})=Y_i^{(\mu)}-M_i^{(\mu)}$$

gilt und dass jeweils $Y_i^{(\mu)}-M_i^{(\mu)}$ auf $Y_j^{(\mu)}-M_j^{(\mu)}$ $(j_1 \leqq j \leqq i)$ vermöge α_{ij} biholomorph bezogen wird. $\check{M}^{(\mu)}$und $M_i^{(\mu)}$ sind entweder leer oder höchstens $(n-1)$-dimensionale komplexe Räume. Ist $\check{U}^* \cap \check{M}^{(\mu)}$ (also auch $U_i^* \cap M_i^{(\mu)}$) nichtleer, so gibt es auf Grund der Induktionsvoraussetzung einen Index $j_0^{(\mu)}$ mit $j_1 \leqq j_0^{(\mu)}$, derart, dass für $j_0^{(\mu)} \leqq j \leqq i$ stets $U_i^* \cap M_i^{(\mu)}$ auf $U_j^* \cap M_j^{(\mu)}$ vermöge der Beschränkung von α_{ij} biholomorph, also insbesondere bijektiv abgebildet wird; ist $\check{U}^* \cap \check{M}^{(\mu)}$ (also auch $U_i^* \cap M_i^{(\mu)}$) leer, so sei $j_0^{(\mu)}:=j_1$. Sei nun j_0 so gewählt, dass $j_0^{(\mu)} \leqq j_0$ für $\mu=1,\cdots,m$. Die α_{ij}^* sind dann für $j_0 \leqq j \leqq i$ bijektiv. Sie sind auch topologisch, da sie als nirgends entartete holomorphe Abbildungen insbesondere offen sind; mithin sind U_i^* und U_j^* homöomorph, falls $j_0 \leqq j \leqq i$.

Sei $('U, \nu, U_{j_0}^*)$ eine Normalisierung des komplexen Raumes $U_{j_0}^*$; mit $'\mathfrak{O}$ bezw. $\mathfrak{O}_i^*$ seien die Strukturgarben von $'U$ bezw. U_i^* bezeichnet. $'\mathfrak{O}$ und die $\mathfrak{O}_i^*$ mit $j_0 \leqq i$ können als analytische Garben über $U_{j_0}^*$ aufgefasst werden, sodass $\mathfrak{O}_j^* \subseteq \mathfrak{O}_i^* \subseteq {}'\mathfrak{O}$ für $j_0 \leqq j \leqq i$ gilt; sie sind als solche kohärent (siehe [7, 18], sowie [12], Satz 27). U_{j_0} ist ein relativ kompakter offener Teilraum von $U_{j_0}^*$. Wäre nun die Behauptung nicht richtig, so würde es eine abzählbar unendliche total-geordnete Teilmenge $i_1 \leqq i_2 \leqq \cdots \leqq i_n \leqq \cdots$ von $\mathfrak{J}$ geben, derart, dass stets $\mathfrak{O}_{i_n}^* \mid U_{j_0}$ in $\mathfrak{O}_{i_{n+1}}^* \mid U_{j_0}$ echt enthalten ist. Dies stände aber in Widerspruch zu einem Satz von H. Grauert ([10], §2, Satz 8). Also existiert ein $i_0 \in \mathfrak{J}$, sodass $'\alpha_{ij}: U_i \to U_j$ für $i_0 \leqq j \leqq i$ biholomorph ist, w. z. b. w.

Wir beweisen nun

SATZ I. *Sei $f: X \to Y$ eine holomorphe Abbildung des irreduziblen komplexen Raumes X in den komplexen Raum Y. Es existiere eine analy-*

tische Menge A in X, derart, dass die Abbildung $f \mid A: A \to Y$ *eigentlich ist und dass* $r(f \mid A) = r(f)$ *gilt. Dann gibt es eine mit f im engeren Sinne verwandte maximale holomorphe Abbildung* $f^*: X \to Y^*$ *auf einen* $r(f)$*-dimensionalen komplexen Raum* Y^*.

Das Paar (f^*, Y^*) ist also eine komplexe Basis zur Abbildung f.

Beweis. A darf als irreduzibel vorausgesetzt werden. Weiter darf angenommen werden, dass f surjektiv und dass Y irreduzibel und $r(f)$-dimensional ist. Trifft dies noch nicht zu, so ist nach Proposition 1.2.2 jedenfalls $f(A) = f(X) := {}'Y$ eine $r(f)$-dimensionale irreduzible analytische Menge in Y. Sei $'f: X \to {}'Y$ die durch Beschränkung von f bestimmte holomorphe Abbildung; $'f$ ist mit f im engeren Sinne verwandt, hat die für f vorausgesetzten Eigenschaften und ist surjektiv. f und Y können jetzt durch $'f$ und $'Y$ ersetzt werden.—Es mögen weiter die alten Bezeichnungen gelten.

Wir betrachten nun Paare (f_i, Y_i), wo Y_i ein irreduzibler $r(f)$-dimensionaler komplexer Raum ist und $f_i: X \to Y_i$ eine mit f im engeren Sinne verwandte surjektive holomorphe Abbildung, die f majorisiert. Zwei Paare (f_i, Y_i), (f_j, Y_j) mögen β-äquivalent heissen, wenn es eine biholomorphe Abbildung $\beta_{ij}: Y_i \to Y_j$ gibt, sodass $f_j = \beta_{ij} \circ f_i$ ist. Die Gesamtheit der β-Äquivalenzklassen ist eine sinnvoll definierte nichtleere Menge. Sind $\mathfrak{b}_0$, $\mathfrak{b}_0'$ zwei β-Äquivalenzklassen, derart, dass für zwei Repräsentanten (f_0, Y_0), (f_0', Y_0') von $\mathfrak{b}_0$ und $\mathfrak{b}_0'$ gilt, dass f_0 von f_0' majorisiert wird, so schreiben wir $\mathfrak{b}_0 \leqq \mathfrak{b}_0'$. Die Relation $\leqq$ ist eine Ordnungsrelation; die auf diese Weise geordnete Menge der β-Äquivalenzklassen werde mit $\mathfrak{B}$ bezeichnet.

$\mathfrak{B}$ ist vermöge der Relation $\leqq$ eine gerichtete Menge: Sind $\mathfrak{b}_1$, $\mathfrak{b}_2$ irgend zwei β-Äquivalenzklassen und (f_1, Y_1), (f_2, Y_2) Repräsentanten von $\mathfrak{b}_1$, $\mathfrak{b}_2$, so ist die Verbindung $(f_1, f_2): X \to Y_1 \times Y_2$ eine mit f im engeren Sinne verwandte holomorphe Abbildung; es ist also $r((f_1, f_2)) = r(f) = r(f \mid A) = r((f_1, f_2) \mid A)$. Ferner ist $(f_1, f_2) \mid A = (f_1 \mid A, f_2 \mid A)$ mit $f_1 \mid A$ und $f_2 \mid A$ eigentlich. Nach Proposition 1.2.2 ist demnach $(f_1, f_2)(A) = (f_1, f_2)(X) =: Y_3$ eine irreduzible $r(f)$-dimensionale analytische Menge in $Y_1 \times Y_2$. Es bezeichne $f_3: X \to Y_3$ die durch Beschränkung von (f_1, f_2) bestimmte holomorphe (surjektive) Abbildung; f_3 ist mit f im engeren Sinne verwandt unde majorisiert f_1 und f_2 (also auch f). Für die durch (f_3, Y_3) repräsentierte β—Äquivalenzklasse $\mathfrak{b}_3$ gilt daher $\mathfrak{b}_1 \leqq \mathfrak{b}_3$ und $\mathfrak{b}_2 \leqq \mathfrak{b}_3$.

Wir repräsentieren nun die Klassen aus $\mathfrak{B}$ durch Paare (f_i, Y_i), wobei jetzt i eine zu $\mathfrak{B}$ isomorph geordnete gerichtete Menge $\mathfrak{J}$ durchläuft, Zu je zwei Paaren (f_i, Y_i), (f_j, Y_j) mit $i, j \in \mathfrak{J}$ und $j \leqq i$ existieren dann (surjektive) holomorphe Abbildungen $\alpha_i: Y_j \to Y$, $\alpha_j: Y_j \to Y$, $\alpha_{ij}: Y_i \to Y_j$ mit

$f = \alpha_i \circ f_i = \alpha_j \circ f_j$ und $\alpha_i = \alpha_j \circ \alpha_{ij}$. Es ist auch $f \mid A = \alpha_i \circ f_i \mid A$, daher ist $f_i \mid A$ stets eigentlich. Wegen der Verwandtschaft von f und f_i ist weiter $r(f) = r(f_i)$ und $r(f \mid A) = r(f_i \mid A)$, also $r(f_i) = r(f_i \mid A)$. Nach Proposition 1.2.2 ist mithin $f_i(A) = f_i(X) = Y_i$, demnach ist $f_i \mid A$ surjektiv. woraus folgt—wiederum wegen $f \mid A = \alpha_i \circ f_i \mid A$—dass α_i eigentlich ist.

Für jeden Punkt $y \in Y$ ist die Faser $F_y := (f \mid A)^{-1}(y)$ eine kompakte analytische Menge, die in endlich viele, etwa in $c(y)$, zusammenhängende Komponenten zerfällt. Es folgt, dass für jedes $i \in \mathfrak{J}$ die Menge $\alpha_i^{-1}(y) \subset Y_i$ endlich ist: Wegen $\alpha_i^{-1}(y) = f_i(F_y)$ besteht sie aus höchstens $c(y)$ Punkten. α_i ist mithin eine eigentliche Überlagerungsabbildung und das Tripel (Y_i, α_i, Y) eine analytische Überlagerung von Y, etwa mit der Blätterzahl b_i. Wegen $\alpha_i = \alpha_j \circ \alpha_{ij}$ $(j \leqq i)$ sind dann auch die α_{ij} eigentliche Überlagerungsabbildungen und die Tripel (Y_i, α_{ij}, Y_j) analytische Überlagerungen.

Die Blätterzahlen b_i müssen sämtlich unter einer endlichen Schranke bleiben. Sei nämlich M_l für die natürliche Zahl l die Menge derjenigen Punkte $y \in Y$, für welche $c(y) \leqq l$ gilt. Nach Proposition 1.2.3 gibt es dann eine natürliche Zahl l_0 und eine nichtleere offene Menge E in Y, derart, dass M_{l_0} in E dicht liegt. Andererseits existiert zur Überlagerung (Y_i, α_i, Y) eine in Y dichte offene Menge E_i, sodass über jedem Punkte von E_i vermöge α_i genau b_i Punkte von Y_i liegen. Der Durchschnitt $M_{l_0} \cap E_i$ ist nicht leer. Für jeden Punkt $\bar{y} \in M_{l_0} \cap E_i$ gilt dann $b_i \leqq c(\bar{y})$ und $c(\bar{y}) \leqq l_0$, mithin ist stets $b_i \leqq l_0$.

Sei $m \in \mathfrak{J}$ ein Index, sodass b_m maximal ist; für $m \leqq j$ ist dann stets (Y_j, α_{jm}, Y_m) eine analytische Überlagerung der Blätterzahl 1. Es sei für $m \leqq j$ jeweils $('Y_j, \nu_j, Y_j)$ eine Normalisierung von Y_j. Die Abbildung $\alpha_{jm}: Y_j \to Y_m$ lässt sich zu einer holomorphen Abbildung $'\alpha_{jm}: 'Y_j \to 'Y_m$ liften, sodass $\alpha_{jm} \circ \nu_j = \nu_m \circ '\alpha_{jm}$ gilt (vgl. [26]); dabei ist $'\alpha_{jm}$ wieder eine eigentliche Überlagerungsabbildung. Da auch die Blätterzahlen der analytischen Überlagerungen $('Y_j, \nu_j, Y_j)$ sämtlich 1 sind, ist $('Y_j, '\alpha_{jm}, 'Y_m)$ stets eine analytische Überlagerung der Blätterzahl 1. Dann aber muss, da $'Y_j$ und $'Y_m$ normale komplexe Räume sind, $'\alpha_{jm}$ biholomorph sein. Infolgedessen ist auch $('Y_m, \nu_j \circ '\alpha_{jm}^{-1}, Y_j)$ eine Normalisierung von Y_j. Es sei $\phi_j := \nu_j \circ '\alpha_{jm}^{-1}$ gesetzt. Ist weiter $j_1 \in \mathfrak{J}$ beliebig, so gibt es ein $j \in \mathfrak{J}$ mit $m \leqq j$ und $j_1 \leqq j$; sei $\phi_{j_1} := \alpha_{jj_1} \circ \phi_j$. Für das System der komplexen Räume Y_i, $\check{Y} := 'Y_m$ und der Abbildungen α_{ij}, ϕ_i sind dann die in der Voraussetzung von Proposition 1.3.1 angegebenen Forderungen erfüllt.

Es sei jetzt Y^* der (hier zunächst nur als Menge betrachtete) inverse Limes des durch die Y_i, α_{ij} bestimmten inversen Systems: Die Elemente von Y^* sind geordnete Punktmengen $y^* := \{\cdots, y_i, \cdots\}$, $i \in \mathfrak{J}$, wobei $y_i \in Y_i$

und $\alpha_{ij}(y_i) = y_j$ für $j \leqq i$ ist. Es seien $f^*: X \to Y^*$ bezw. $\psi_i: Y^* \to Y_i$ die durch die Zuordnungen

$$x \to \{\cdot\cdot\cdot, f_i(x), \cdot\cdot\cdot\}, \; x \in X, \text{ bezw. } y^* = \{\cdot\cdot\cdot, y_i, \cdot\cdot\cdot\} \to y_i$$

gegebenen Abbildungen; man hat $f_i = \psi_i \circ f^*$ und $\psi_j = \alpha_{ij} \circ \psi_i$, falls $j \leqq i$. Mit Hilfe von Proposition 1.3.1 ergibt sich nun: Zu jedem Element $y^* = \{\cdot\cdot\cdot, y_i, \cdot\cdot\cdot\} \in Y^*$ existiert jeweils in Y_i eine relativ kompakte offene Umgebung $U_i(y_i; y^*)$ von y_i mit $\alpha_{ij}^{-1}(U_j(y_j; y^*)) = U_i(y_i; y^*)$ für $j \leqq i$, sowie ein Index $i_0(y^*) = i_0$, sodass für $i_0 \leqq j \leqq i$ stets $U_i(y_i; y^*)$ auf $U_j(y_j; y^*)$ vermöge α_{ij} biholomorph bezogen wird. Es folgt insbesondere, dass $y^* \in f^*(X)$, dass also f^* surjektiv ist. Weiter folgt, dass die Menge $\psi_i^{-1}(U_i(y_i; y^*))$ nicht vom Index i abhängt—sie sei mit $V(y^*)$ bezeichnet—und dass die durch Beschränkung von ψ_i bestimmte Abbildung $'\psi_i: V(y^*) \to U_i(y_i; y^*)$ für $i_0(y^*) \leqq i$ bijektiv ist. Vermöge der Abbildung $'\psi_{i_0}^{-1}$ werde nunmehr die Topologie und komplexe Struktur von $U_{i_0}(y_{i_0}; y^*)$ auf $V(y^*)$ übertragen. Insgesamt ist so, wenn y^* die Elemente von Y^* durchläuft, auf Y^* eine Hausdorffsche Topologie und eine mit ihr verträgliche komplexe Struktur festgelegt. Y^* wird auf diese Weise zu einem rein $r(f)$-dimensionalen komplexen Raum; die Abbildung $f^*: X \to Y^*$ wird zu einer mit f im engeren Sinne verwandten (surjektiven) holomorphen Abbildung, und die $\psi_i: Y^* \to Y_i$ werden zu nirgends entarteten holomorphen Abbildungen. Wegen $f_i = \psi_i \circ f^*$ werden alle f_i, also auch f, von f^* majorisiert; das Paar (f^*, Y^*) repräsentiert daher ein Element von $\mathfrak{B}$.

Wir behaupten, dass f^* eine gesuchte maximale Abbildung ist. Hierzu braucht nur noch gezeigt zu werden, dass f^* jede von f strikt abhängige holomorphe Abbildung majorisiert. Sei $\phi_a: X \to Y_a$ eine solche Abbildung. Die Verbindung $\Phi := (f^*, \phi_a): X \to Y^* \times Y_a$ ist mit f^*, also mit f, im engeren Sinne verwandt, es ist daher $r(\Phi) = r(f) = r(f \mid A) = r(\Phi \mid A)$; ferner ist $\Phi \mid A$ mit $f^* \mid A$ eigentlich. Nach Proposition 1.2.2 ist demnach $\Phi(A) = \Phi(X) =: Y'$ eine irreduzible $r(f)$-dimensionale analytische Menge in $Y^* \times Y_a$. Es bezeichne $\epsilon: Y' \to Y$ die Injektion von Y' in Y und $\Phi': X \to Y'$ die durch Beschränkung von Φ bestimmte (surjektive) holomorphe Abbildung; ferner seien σ^* und σ_a die Projektionen von $Y^* \times Y_a$ auf Y^* bezw. Y_a. Man hat $f^* = \sigma^* \circ \epsilon \circ \Phi'$, also wird f^* von Φ' majorisiert. Es wird dann auch f von Φ' majorisiert; dass Paar (Φ', Y') repräsentiert mithin ein Element von $\mathfrak{B}$, und daher wird Φ' von f^* majorisiert. Es folgt, dass die Abbildung $\sigma^* \circ \epsilon: Y' \to Y^*$ biholomorph ist. Es ist $\Phi' = (\sigma^* \circ \epsilon)^{-1} \circ f^*$ und weiter $\phi_a = \sigma_a \circ \epsilon \circ \Phi'$, also gilt $\phi_a = (\sigma_a \circ \epsilon \circ (\sigma^* \circ \epsilon)^{-1}) \circ f^*$. Demnach wird, wie behauptet, ϕ_a von f^* majorisiert.

KOROLLAR 1 zu Satz I. *Ist X normal, so auch Y^*.*

In der Tat: Sei $('Y^*, \nu^*, Y^*)$ eine Normalisierung von Y^*; dann lässt sich $f^*: X \to Y^*$ zu einer holomorphen Abbildung $'f^*: X \to 'Y^*$ liften, sodass $f^* = \nu^* \circ 'f^*$ gilt. $'f^*$ ist mit f^*, also mit f, im engeren Sinne verwandt. Wegen der Maximalität von f^* gibt es eine holomorphe Abbildung $\mu^*: Y^* \to 'Y^*$ mit $'f^* = \mu^* \circ f^*$. Es ist demnach $f^* = \nu^* \circ \mu^* \circ f^*$, woraus folgt, dass $\nu^* \circ \mu^*$: $Y^* \to Y^*$ die Identität, also ν^* biholomorph ist. Y^* ist daher normal.

Es ist naheliegend, wie Satz I zu modifizieren ist, wenn die Voraussetzung der Irreduzibilität für den Raum X fallengelassen wird: Man hat dann für jede irreduzible Komponente X' von X die Übereinstimmung der globalen Ränge von $f \mid X'$ und $f \mid (A \cap X')$ zu fordern. Der Beweis wird nur technisch komplizierter, er lässt sich im Prinzip aber ähnlich wie oben führen.

Die analytische Menge A kann insbesondere mit X zusammenfallen. So ergibt sich das

KOROLLAR 2 zu Satz I. *Zu jeder eigentlichen holomorphen Abbildung $f: X \to Y$ existiert eine mit f im engeren Sinne verwandte maximale eigentliche holomorphe Abbildung $f^*: X \to Y^*$.*

Wir schliessen noch einige Bemerkungen an.

1) Die Aussage des Korollars 2 ergibt sich auch aus dem folgenden Satz, den H. Cartan in [8] unter Verwendung des Hauptresultates von H. Grauert in [10] über die Kohärenz von Bildern analytischer Garben bewiesen hat: *Sei $f: X \to Y$ eine holomorphe Abbildung, derart, dass die Niveaumengen von f sämtlich kompakt sind; sei R die durch diese Niveaumengen definierte Äquivalenzrelation in X. Dann ist die auf dem Quotientenraum X/R definierte geringte Quotientenstruktur eine komplexe Struktur.* Entsprechende Resultate für den Fall, dass X eine komplexe Mannigfaltigkeit ist, und in einen anderen Spezialfall wurden schon in [29] gewonnen. Aus dem zitierten Satz von H. Cartan folgt ausserdem, dass eine eigentliche maximale holomorphe Abbildung stets *einfach* ist, d. h., dass ihre Fasern zusammenhängend sind.

2) Es ist möglich, Satz I wie folgt auszudehnen: Zu der holomorphen Abbildung $f: X \to Y$ werden spezielle Klassen von f strikt abhängiger holomorpher Abbildungen betrachtet, und es wird gefragt, ob es in einer derartigen Klasse eine Abbildung gibt, die alle Abbildungen der Klasse majorisiert. Im Rahmen solcher Betrachtungen lässt sich dann auch ein Analogon zu Satz I für holomorphe Abbildungskeime auf komplexen Unterräumen eines komplexen Raumes gewinnen. Wir verzichten hier darauf, dies durchzuführen. Im Paragraphen 5 werden aber die entsprechenden Fragen für den Fall

meromorpher Abbildungen erörtert. Der dort benutzte Begriff der Abhängigkeit meromorpher Abbildungen deckt sich im Spezialfall holomorpher Abbildungen mit der oben erklärten Abhängigkeit im weiteren Sinne. Daher wird mit den Ergebnissen des Paragraphen 5 insbesondere auch eine Antwort auf die Frage gegeben, was an die Stelle der Resultate dieses Abschnittes 1.3 zu treten hat, wenn der bisher für holomorphe Abbildungen ausschliesslich benutzte Begriff der strikten Abhängigkeit ersetzt wird durch den allgemeineren Begriff der Abhängigkeit schlechthin.

2. Eigenschaften maximaler holomorpher Abbildungen. Besondere Fälle. Die in den Betrachtungen dieses Paragraphen auftretenden komplexen Räume X, Y werden der Einfachheit halber stets als irreduzibel vorausgesetzt.

2.1. Wir diskutieren zunächst eine topologische Eigenschaft maximaler holomorpher Abbildungen.

Jede stetige Abbildung $f: X \to Y$ induziert einen Homomorphismus $\mathring{f}: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ der Fundamentalgruppen ($x_0 \in X$ sei ein beliebig gewählter Bezugspunkt). Es gilt

PROPOSITION 2.1.1. *Ist $f: X \to Y$ eine maximale holomorphe Abbildung, so ist der Homomorphismus $\mathring{f}: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ ein Epimorphismus.*

Beweis. Zur Untergruppe $\mathring{f}(\pi_1(X, x_0)) =: \mathfrak{U}$ von $\pi_1(Y, f(x_0))$ gibt es eine unverzweigte unbegrenzte Überlagerung (Y', τ, Y) von Y, derart, dass für einen Punkt $y_0' \in Y'$ mit $\tau(y_0') = f(x_0) =: y_0$ die Projektionsabbildung $\tau: Y' \to Y$ einen Isomorphismus von $\pi_1(Y', y_0')$ auf $\mathfrak{U}$ induziert. Y' kann mit einer komplexen Struktur versehen werden und wird dementsprechend als ein komplexer Raum aufgefasst sodass τ eine lokal-biholomorphe Abbildung wird. Die Abbildung $f: X \to Y$ lässt sich wie folgt zu einer Abbildung $f': X \to Y'$ liften. Ist $x \in X$ irgendein Punkt, so sei $\mathfrak{C}(x_0, x)$ ein x_0 mit x verbindendes Kurvenstück. Vermöge f geht $\mathfrak{C}(x_0, x)$ in ein Kurvenstück $\mathfrak{C}_1$ in Y mit $y_0 = f(x_0)$ als Anfangspunkt über; $\mathfrak{C}_1$ kann zu einem Kurvenstück $\mathfrak{C}_1'$ in Y' mit y_0' als Anfangspunkt geliftet werden. Bezeichnet y' den Endpunkt von $\mathfrak{C}_1'$, so werde jetzt $f'(x) := y'$ gesetzt. Es folgt auf Grund der Festlegung von Y' leicht, dass y' nur von x, nicht aber von der Wahl von $\mathfrak{C}(x_0, x)$ abhängt. Demnach ist $f': X \to Y'$ eindeutig definiert, und es gilt $f = \tau \circ f'$. Da τ lokal-biholomorph ist, ist f' eine holomorphe und mit f im engeren Sinne verwandte Abbildung. Wegen der Maximalität von f existiert nun eine holomorphe Abbildung $\sigma: Y \to Y'$ mit $f' = \sigma \circ f$; man hat daher

$f = \tau \circ \sigma \circ f$. Hieraus ergibt sich, dass σ injektiv und lokal biholomorph abbildet; infolgedessen ist $f'(X) = \sigma(Y) = : Y''$ ein offener Teilraum von Y', der vermöge τ homöomorph auf Y bezogen wird. Dann aber muss notwendig Y'' mit Y' zusammenfallen; $\tau: Y' \to Y$ ist also ein Homöomorphismus und somit $\mathfrak{U} = \pi_1(Y, f(x_0))$, w. z. b. w.

Folgerung. Ist die holomorphe Abbildung $f: X \to Y$ maximal und X einfach zusammenhängend, so ist auch Y einfach zusammenhängend.

2.2. Wir gehen in diesem Abschnitt von der folgenden Situation aus: $f: X \to Y$ sei eine holomorphe Abbildung, zu der eine komplexe Basis (f^*, Y^*) existiert. Es sollen Fälle betrachtet werden, in denen der Raum Y^* kompakt ist. In einem solchen Falle ist dann insbesondere jede von f abhängige holomorphe Funktion konstant.

Da eine maximale Abbildung stets surjektiv ist, ist Y^* sicher dann kompakt, wenn X kompakt ist. Allgemeiner folgt die Kompaktheit von Y^* auch dann, wenn in X eine kompakte analytische Menge A existiert, derart, dass wie in Satz I die Abbildung $f \mid A: A \to Y$ den gleichen globalen Rang wie f in X hat. Dann ist auch $r(f^*) = r(f^* \mid A)$; nach Proposition 1.2.2 ist daher $f^*(A) = f^*(X) = Y^*$ und infolgedessen Y^* kompakt.

In manchen Fällen lässt sich zeigen, dass die Dimension von Y^* mit dem globalen Rang von f übereinstimmt. Z. B. trifft dies zu, wenn f die Voraussetzungen von Satz I erfüllt, oder auch, wenn der lokale Rang von f konstant ist (vgl. [30]), ferner, wenn X abzählbare Topologie hat. Ob allgemein $r(f) = \dim Y^*$ gilt, muss offen bleiben; wir wollen es hier für die folgende Aussage voraussetzen.—Es sei $\psi: Y^* \to Y$ die holomorphe Abbildung, für welche $f = \psi \circ f^*$ gilt. Für jede natürliche Zahl k seien weiter mit $H_k(X)$, $H_k(Y)$, $H_k(Y^*)$ die k-ten ganzzahligen Homologiegruppen von X, Y, Y^* bezeichnet. Man hat durch f, f^*, ψ induzierte Homomorphismen

$$\bar{f}_k: H_k(X) \to H_k(Y), \ \bar{f}_k^*: H_k(X) \to H_k(Y^*), \ \bar{\psi}_k: H_k(Y^*) \to H_k(Y)$$

mit $\bar{f}_k = \bar{\psi}_k \circ \bar{f}_k^*$. Es gilt nun:

Sei $r := r(f)$ der globale Rang von f in X. Ist dann $\bar{f}_{2r}: H_{2r}(X) \to H_{2r}(Y)$ nicht der Nullhomomorphismus, so ist Y^* kompakt.

Zum Beweise hat man zu benutzen, dass für einen irreduziblen komplexen Raum der Dimension n die $2n$-te ganzzahlige Homologiegruppe genau dann die Nullgruppe ist, wenn der Raum nicht kompakt ist ([4]). Im vorliegenden Fall ist $n = \dim Y^* = r$. Wäre nun Y^* nicht kompakt, so wäre $H_{2r}(Y^*) = 0$.

Wegen $\bar{f}_{2r} = \bar{\psi}_{2r} \circ \bar{f}_{2r}^*$ müsste aber dann $\bar{f}_{2r}$ notwendig der Nullhomomorphismus sein, entgegen der Voraussetzung.

Es sei nun weiter X als normal vorausgesetzt, und Y sei die Riemannsche Zahlenkugel $\boldsymbol{P}^1$; die holomorphe Abbildung $f\colon X \to \boldsymbol{P}^1$ ist also eine meromorphe Funktion ohne Unbestimmtheitsstellen. $f^*\colon X \to Y^*$ existiert dann stets, und Y^* ist eine Riemannsche Fläche. Wir behaupten: *Gestattet f keine Darstellung als Quotient teilerfremder holomorpher Funktionen in X, so ist Y^** kompakt. In der Tat, wäre Y^* nicht kompakt, so liesse die auf Y^* meromorphe Funktion $\psi\colon Y^* \to \boldsymbol{P}^1$, für welche $f = \psi \circ f^*$ gilt, eine teilerfremde Quotientendarstellung $\psi = \psi_0/\psi_1$ mit holomorphen Funktionen ψ_0, ψ_1 zu; dann aber hätte man in $f = \psi_0 \circ f^*/\psi_1 \circ f^*$ eine Darstellung von f als Quotient der in X teilerfremden holomorphen Funktionen $\psi_0 \circ f^*$ und $\psi_1 \circ f^*$.— Es folgt, dass unter der angegebenen Voraussetzung über f jede von f abhängige meromorphe Funktion (die dann ebenfalls keine Unbestimmtheitsstellen aufweist) von f algebraisch abhängig ist (vgl. hierzu auch § 6 in Teil II dieser Arbeit).

Eine meromorphe Funktion *mit* Unbestimmtheitsstellen definiert lediglich eine *meromorphe* Abbildung in den $\boldsymbol{P}^1$, zu der eine maximale meromorphe Abbildung auf eine kompakte Riemannische Fläche gehört, wie sich aus den Resultaten der folgenden Paragraphen ergeben wird.

2.3. Wie in 2.2 sei $f\colon X \to Y$ eine holomorphe Abbildung, zu der eine komplexe Basis (f^*, Y^*) existiert. Es sei wieder $\psi\colon Y^* \to Y$ die holomorphe Abbildung mit $f = \psi \circ f^*$; ψ ist nirgends entartet. Mit $\mathfrak{R}(X;f)$ werde der Ring der von f abhängigen holomorphen Funktionen auf X bezeichnet. Wegen der Maximalität von $f^*\colon X \to Y^*$ existiert zu jeder Funktion $\phi \in \mathfrak{R}(X;f)$ eine auf Y^* holomorphe Funktion ϕ^* mit $\phi = \phi^* \circ f$, und umgekehrt bestimmt jede auf Y^* holomorphe Funktion ϕ^* eine holomorphe Funktion $\phi := \phi^* \circ f \in \mathfrak{R}(X;f)$. Somit induziert f^* eine bijektive Abbildung ${}^*f^*\colon \mathfrak{R}(Y^*) \to \mathfrak{R}(X;f)$ des Ringes der auf Y^* holomorphen Funktionen auf den Ring $\mathfrak{R}(X;f)$, und diese Abbildung ist ein Isomorphismus.

Ist nun ψ surjektiv und eigentlich, so ist (Y^*, ψ, Y) eine analytische Überlagerung von Y, etwa der Blätterzahl b. Weiter mögen X und Y als normal angenommen werden (woraus folgt, dass Y^* ebenfalls normal ist). Der Ring $\mathfrak{R}(Y^*)$ kann dann als ein ganzer Ring endlichen Grades $[\mathfrak{R}(Y^*) : \mathfrak{R}(Y)]$ über dem Ring $\mathfrak{R}(Y)$ der auf Y holomorphen Funktionen aufgefasst werden mit $[\mathfrak{R}(Y^*) : \mathfrak{R}(Y)] \leqq b$. ψ ist insbesondere dann surjektiv und eigentlich, wenn f surjektiv ist und ausserdem die Voraussetzungen von Satz I erfüllt sind. Demnach gilt

PROPOSITION 2.3.1. *Sei $f: X \to Y$ eine holomorphe Abbildung des irreduziblen normalen komplexen Raumes X auf den normalen komplexen Raum Y. Es existiere eine analytische Menge A in X, derart, dass die Abbildung $f \mid A: A \to Y$ eigentlich ist und dass $r(f \mid A) = r(f)$ gilt. Dann ist $\mathfrak{R}(X;f)$ zu einem ganzen Ring endlichen Grades über $\mathfrak{R}(Y)$ isomorph.*

2.4. Es soll weiter auf eine Eigenschaft besonderer maximaler Abbildungen von Holomorphiehüllen komplexer Räume eingegangen werden. Die auftretenden komplexen Räume werden sämtlich als irreduzibel und normal vorausgesetzt.

Sei $\tau: {}'X \to X$ eine nirgends entartete holomorphe Abbildung, es sei $\dim X = \dim {}'X$; das Tripel $('X, \tau, X)$ heisst ein *Gebiet über* X. Bezeichnet $\mathfrak{A}$ eine Menge holomorpher Abbildungen $\alpha: {}'X \to Z_\alpha$, von $'X$ in komplexe Räume Z_α, so lässt sich $('X, \tau, X)$ eine *Holomorphiehülle* $H(('X, \tau, X); \mathfrak{A})$ zuordnen (vgl. [14]), und zwar ist $H(('X, \tau, X); \mathfrak{A})$ ein Gebiet $(''X, \sigma, X)$ über X mit folgenden Eigenschaften: a) Es gibt eine nirgends entartete holomorphe Abbildung $\mu: {}'X \to {}''X$ mit $\tau = \sigma \circ \mu$, derart, dass jede Abbildung $\alpha \in \mathfrak{A}$ vermöge μ nach $''X$ holomorph fortsetzbar ist (d.h. zu $\alpha: {}'X \to Z_\alpha$ existiert eine holomorphe Abbildung $\beta: {}''X \to Z_\alpha$ mit $\alpha = \beta \circ \mu$). b) Ist $('Y, \xi, X)$ ein Gebiet über X und gibt es eine nirgends entartete holomorphe Abbildung $\eta: {}'X \to {}'Y$ mit $\tau = \xi \circ \eta$, sodass jede Abbildung $\alpha \in \mathfrak{A}$ vermöge η nach $'Y$ fortsetzbar ist, so existiert eine holomorphe Abbildung $\psi: {}'Y \to {}''X$ mit $\mu = \psi \circ \eta$.—Ist insbesondere $\sigma: {}''X \to X$ biholomorph, so kann $''X$ mit X identifiziert werden; wir schreiben dann $H(('X, \tau, X); \mathfrak{A}) = X$.

Es sei nun eine maximale holomorphe Abbildung $f: X \to Y$ von konstantem lokalen Rang gegeben. Die holomorphe Abbildung $'f: = f \circ \tau: {}'X \to Y$ hat dann ebenfalls konstanten lokalen Rang, mithin existiert zu $'f$ eine komplexe Basis $('f^*, Y^*)$ (vgl. [30]). Es ist also $'f^*: {}'X \to Y^*$ eine mit $'f$ im engeren Sinne verwandte maximale holomorphe Abbildung, und man hat eine holomorphe und insbesondere nirgends entartete Abbildung $\tau^*: Y^* \to Y$ mit $'f = \tau^* \circ {}'f^*$. Demnach ist (Y^*, τ^*, Y) ein Gebiet über Y. Sei jetzt weiter $\mathfrak{A}^*$ eine Menge von holomorphen Abbildung $\alpha^*: Y^* \to Z_{\alpha^*}$, derart, dass alle Abbildungen $\alpha^* \circ {}'f^*: {}'X \to Z_{\alpha^*}$ zu $\mathfrak{A}$ gehören. (Besteht z.B. $\mathfrak{A}$ aus allen holomorphen Abbildungen von $'X$ in einen festen komplexen Raum Z, so ist die Menge aller holomorphen Abbildungen von Y^* in Z eine zulässige Menge $\mathfrak{A}^*$.) Wir behaupten:

Ist $H(('X, \tau, X); \mathfrak{A}) = X$, so ist $H((Y^*, \tau^*, Y); \mathfrak{A}^*) = Y$.

Zum Beweise genügt es zu zeigen, dass jede holomorphe Abbildung $\alpha^*: Y^* \to Z_{\alpha^*}$ aus $\mathfrak{A}^*$ vermöge τ^* nach Y holomorph fortsetzbar ist. Die Abbildung $\alpha^* \circ 'f^*: 'X \to Z_{\alpha^*}$ gehört zu $\mathfrak{A}$, daher ist sie vermöge τ zu einer holomorphen Abbildung $\alpha_0: X \to Z_{\alpha^*}$ fortsetzbar; es gilt also $\alpha^* \circ 'f^* = \alpha_0 \circ \tau$. Die Abbildung $\alpha^* \circ 'f^*$ ist von $'f^*$, also auch von $'f$ abhängig; es folgt, dass in der offenen Menge $\tau('X) \subset X$ die Fortsetzung α_0 von $\alpha^* \circ 'f^*$ von der Fortsetzung f von $'f$ abhängig ist. Da f konstanten lokalen Rang hat, muss dann α_0 in ganz X von f abhängig sein. Wegen der Maximalität von f existiert somit eine holomorphe Abbildung $\alpha_0^*: Y \to Z_{\alpha^*}$ mit $\alpha_0 = \alpha_0^* \circ f$. Man hat $\alpha^* \circ 'f^* = \alpha_0 \circ \tau = \alpha_0^* \circ f \circ \tau$, wegen $f \circ \tau = 'f = \tau^* \circ 'f^*$ also $\alpha^* \circ 'f^* = \alpha_0^* \circ \tau^* \circ 'f^*$, und hieraus ergibt sich $\alpha^* = \alpha_0^* \circ \tau^*$, da $'f^*$ surjektiv ist. Demnach ist α_0^* eine gesuchte Fortsetzung von α^* vermöge τ^* nach Y.

Wir wollen noch den besonderen Fall betrachten, dass Y eine nichtkompakte Riemannsche Fläche ist (der lokale Rang von $f: X \to Y$ ist also 1), dass ferner $\mathfrak{A}$ mit dem Ring $\mathfrak{R}('X)$ der in $'X$ holomorphen Funktionen und $\mathfrak{A}^*$ mit dem Ring $\mathfrak{R}(Y^*)$ der in Y^* holomorphen Funktionen zusammenfällt. Dann ist auch Y^* eine nichtkompakte Riemannsche Fläche und somit eine holomorph vollständige komplexe Mannigfaltigkeit. Es muss also

$$H((Y^*, \tau^*, Y); \mathfrak{R}(Y^*)) = (Y^*, \tau^*, Y)$$

gelten, während unsere obige Aussage lehrt, dass $H((Y^*, \tau^*, Y); \mathfrak{R}(Y^*)) = Y$ ist. Demnach ist τ^* surjektiv und biholomorph, und Y^* kann mit Y identifiziert werden. Es ergibt sich:

Ist $f: X \to Y$ eine maximale holomorphe Abbildung von X auf eine nichtkompakte Riemannsche Fläche, ferner $('X, \tau, X)$ ein Gebiet über X mit $H(('X, \tau, X); \mathfrak{R}('X)) = X$, so ist auch die holomorphe Abbildung $f \circ \tau: 'X \to Y$ maximal.

MÜNCHEN.

LITERATUR.

[1] A. Andreotti and W. Stoll, "Extension of holomorphic maps," *Annals of Mathematics*, vol. 72 (1960), pp. 312-349.

[2] H. Behnke, "Funktionentheorie auf komplexen Mannigfaltigkeiten," *Proceedings of the International Congress of Mathematicians* 1954, vol. III, pp. 45-57.

[3] ——— and H. Grauert, "Analysis in non-compact complex spaces," in *Analytic Functions*, Princeton, 1960.

[4] A. Borel and A. Haefliger, "La classe d'homologie fondamentale d'un espace analytique," *Bulletin de la Société Mathématique de France*, vol. 89 (1961), pp. 461-513.

[5] N. Bourbaki, "Théorie des ensembles," Fascicule de résultats, Paris, 1951.

[6] H. Cartan, *Séminaire É. N. S.*, 1951/52.

[7] ———, *Séminaire É. N. S.*, 1953/54.

[8] ———, "Quotients of complex analytic spaces," *International Colloquium on Function Theory, Tata Institute Bombay*, January 1960, pp. 1-15.

[9] W. L. Chow, "On compact analytic varieties," *American Journal of Mathematics*, vol. 71 (1949), pp. 893-914.

[10] H. Grauert, "Ein Theorem der analytischen Garbentheorie und die Modulräume komplexer Strukturen," *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, No. 5 (1960), pp. 233-292.

[11] ——— and R. Remmert, "Zur Theorie der Modifikationen. I. Stetige und eigentliche Modifikationen komplexer Räume," *Mathematische Annalen*, vol. 129 (1955), pp. 274-296.

[12] ———, "Komplexe Räume," *Mathematische Annalen*, vol. 136 (1958), pp. 245-318.

[13] K. H. Hedtfeld, "Starre einfach-zusammenhängende Holomorphiegebiete," *Schriftenreihe des Mathematischen Instituts der Universität Münster*, Heft 8 (1954).

[14] H. Kerner, "Holomorphiehüllen zu K-vollständigen komplexen Räumen," *Mathematische Annalen*, vol. 138 (1960), pp. 316-328.

[15] K. Koch, "Die analytische Projektion," *Schriftenreihe des Mathematischen Instituts der Universität Münster*, Heft 6 (1953).

[16] N. Kuhlmann, "Zur Theorie der Modifikationen algebraischer Varietäten, *Schriftenreihe des Mathematischen Instituts der Universität Münster*, Heft 14 (1959).

[17] ———, "Projektive Modifikationen komplexer Räume," *Mathematische Annalen*, vol. 139 (1960), pp. 217-238.

[18] ———, "Die Normalisierung komplexer Räume," *Mathematische Annalen*, vol. 144 (1961), pp. 110-125.

[19] K. Oka, "Sur les fonctions analytiques de plusieurs variables. VIII. Lemme fondamental," *Journal of the Mathematical Society of Japan*, vol. 3 (1951), pp. 204-214 and 259-278.

[20] R. Remmert, "Projektionen analytischer Mengen," *Mathematische Annalen*, vol. 130 (1956), pp. 410-441.

[21] ———, "Meromorphe Funktionen in kompakten komplexen Räumen," *Mathematische Annalen*, vol. 132 (1956), pp. 277-288.

[22] ———, "Holomorphe und meromorphe Abbildungen komplexer Räume," *Mathematische Annalen*, vol. 133 (1957), pp. 328-370.

[23] ———, "Sur les espaces analytiques séparables et holomorphiquement convexes," *Comptes Rendus*, vol. 243 (1956), pp. 118-121.

[24] ———, "Analytic and algebraic dependence of meromorphic functions," *American Journal of Mathematics*, vol. 82 (1960), pp. 891-899.

[25] ——— and K. Stein, "Über die wesentlichen Singularitäten analytischer Mengen," *Mathematische Annalen*, vol. 126 (1953), pp. 263-306.

[26] ———, "Eigentliche holomorphe Abbildungen," *Mathematische Zeitschrift*, vol. 73 (1960), pp. 159-189.

[27] J. P. Serre, "Géométrie algébrique et géométrie analytique," *Annales de l'Institut Fourier*, G (1955/56), pp. 1-42.

[28] K. Stein, "Analytische Projektion komplexer Mannigfaltigkeiten," *Colloque sur les fonctions de plusieurs variables*, Bruxelles, 1953, pp. 97-107.

[29] ———, "Analytische Zerlegungen komplexer Räume," *Mathematische Annalen*, vol. 132 (1956), pp. 63-93.

[30] ———, "Die Existenz komplexer Basen zu holomorphen Abbildungen," *Mathematische Annalen*, vol. 136 (1958), pp. 1-8.

[31] W. Stoll, "Über meromorphe Modifikationen I-III," *Mathematische Zeitschrift*, vol. 61 (1954), pp. 204-234, 467-488 and vol. 62 (1955), pp. 189-210.

[32] ———, "Über meromorphe Modifikationen IV, V," *Mathematische Annalen*, vol. 130 (1955), pp. 147-182, 272-316.

[33] ———, "Über meromorphe Abbildungen komplexer Räume I, II," *Mathematische Annalen*, vol. 136 (1958), pp. 201-239, 393-429.

[34] W. Thimm, "Über ausgeartete meromorphe Abbildungen II," *Mathematische Annalen*, vol. 125 (1952), pp. 264-283.

[35] ———, "Über die Menge der singulären Bildpunkte einer meromorphen Abbildung," *Mathematische Zeitschrift*, vol. 57 (1953), pp. 456-480.

[36] ———, "Untersuchungen über ausgeartete meromorphe Abbildungen," *Mathematische Annalen*, vol. 127 (1954), pp. 150-161.

[37] ———, "Meromorphe Abbildungen von Riemannschen Bereichen," *Mathematische Zeitschrift*, vol. 60 (1954), pp. 435-457.

DIFFERENTIABLE PERIODIC ACTIONS ON E^8 WITHOUT FIXED POINTS.*

By J. M. Kister.

Smith has shown [5] that any map T on E^n having a period a power of a prime must have a fixed point. Also, if $n < 7$ and the period of T is a product of two primes then T must have a fixed point providing T is differentiable [6]. Examples were outlined in [3] showing E^{9r} admits a map of period r with no fixed points for each integer r not a power of a prime. These examples will be improved in this paper to show the second result mentioned above is virtually best possible, the single case $n = 7$ being unsettled. More precisely:

THEOREM. *If r is an integer which is not a power of a prime and $n > 7$ then there is a differentiable and simplicial map of period r on E^n without fixed points.*

Proof. If r is not the power of a prime, let $r = pq$ where p and q are relatively prime. Let S_1' and S_2' be 1-spheres which are regular polygons of r sides centered at the origin in the 2-planes E_1^2 and E_2^2, respectively. Define T_1 to be the rotation of E_1^2 through $360/p$ degrees and T_2 to be the rotation of E_2^2 through $360/q$ degrees. Let E^4 be $E_1^2 \times E_2^2$ and T be the orthogonal transformation $T_1 \times T_2$. In E^4 there is a polyhedral 3-sphere denoted by S^3 defined by taking the union of all line segments from a point in S_1' to a point in S_2'. In fact S^3 may be regarded in a natural way as the complex obtained as the join $S_1' \circ S_2'$. Furthermore $T(S^3) = S^3$ and $T \mid S^3$ is simplicial, without fixed points and has period r.

It is shown in 3.1 of [2] how to construct a map $f\colon S^3 \to S^3$ having degree 0 such that $Tf = fT$. Also, denoting the k-th barycentric subdivision of S^3 by $S^3(k)$, f may be assumed simplicial from $S^3(k)$ to S^3 for some k, cf. 5.5 of [2].

Let E_1^4 and E_2^4 be two copies of E^4 containing as subcomplexes $S^3(k) = S_1^3$ and $S^3 = S_2^3$ respectively, and let S^7 denote the 7-sphere in $E^8 = E_1^4 \times E_2^4$ consisting of all line segments joining a point in S_1^3 to a point in S_2^3. Then S^7 has the natural join structure of $S^3(k) \circ S^3$, relative to which $T \times T$ is simplicial.

* Received March 13, 1963.

Next we define an invariant 4-dimensional subcomplex C of S^7 as follows. Order the vertices of S_1^3 by first listing the vertices in $S^3(k-1)$, then the barycenters of 1-simplexes in $S^3(k-1)$, etc., obtaining $V_1, V_2, \cdots, V_m$ and regard f as a map from S_1^3 to S_2^3. C consists of all closed simplexes spanned by $v_{i_0}, \cdots, v_{i_j}, f(v_{i_j}), \cdots, f(v_{i_3})$ where $i_0 < i_1 < i_2 < i_3$ and $0 \leqq j \leqq 3$. Then, since T preserves the ordering on the vertices of any single simplex in S_1^3, and since f commutes with T, C is invariant under $T \times T$ which is simplicial on C. The complex C is a polyhedral form of the mapping cylinder determined by f, and has S_2^3 as a deformation retract. Also, since f is inessential, S_1^3 can be contracted to a point in C, cf. 5.2 in [2].

Finally we subdivide S^7 obtaining the new complex S_*^7 by taking the k-th barycentric subdivision of S_2^3, say $S_2^3(k)$, and letting the simplicial structure on S_*^7 be that of $S_1^3 \circ S_2^3(k)$. This induces a subdivision on C, say C_*, and $T \times T$ is still simplicial on both S_*^7 and C_*.

Let $\rho: E_1^4 \times E_2^4 \to E_1^4 \times E_2^4$ be defined by $\rho(x,y) = (y,x)$ and let $C_*' = \rho(C_*)$, the "reverse" of C_*, another invariant set on which $T \times T$ is simplicial.

We note that S_*^7 bounds a convex 8-cell D in E^8 which we triangulate by taking the cone over S_*^7 from the origin. The next step will be to define an invariant 4-dimensional contractible (infinite) polyhedron K in D-{origin}. Let h_i be the linear map of E^8 defined by $h_i(x) = x/i + 1$. Let S_i be the polyhedron $h_i(S_*^7)$, $i = 0, 1, 2, \cdots$. Since $T \times T$ is orthogonal on E^8, each S_i is invariant, as is $h_i(C_*)$, $h_i(C_*')$, $h_i(S_1^3)$ and $h_i(S_2^3)$. Let I_j^4 be the cone over S_j^3 from the origin, $j = 1, 2$.

We obtain an invariant triangulation of D-{origin} by triangulating each annulus A_i, bounded by S_i and S_{i+1}. This can be accomplished by starring each lateral face of each truncated 8-simplex σ in A_i and then adding a vertex at the center of gravity of σ, without introducing any new vertices on the boundary of A_i.

Then K is the subcomplex in this triangulation consisting of $\cup [h_i(C_*) \cup (A_{i+1} \cap I_1^4)]$, for all i odd, together with $\cup [h_i(C_*') \cup (A_{i-1} \cap I_2^4)]$, for all positive even i. K is invariant under $T \times T$ since it is the union of invariant subcomplexes. $T \times T \mid K$ has period r and is without fixed points. Through successive deformations of the type that takes C onto S_2^3 alternating with those taking the 4-dimensional annuli $A_i \cap I_j^4$ onto $S_{i+1} \cap I_j^4$ it is possible to deform any finite subcomplex of K onto $h_i(S_1^3)$ for large odd i. Since S_1^3 can be contracted to a point in C it follows that any finite subcomplex of K can be contracted to a point in K, hence K is contractible.

Let U be the open star neighborhood of K in the second barycentric

subdivision of the triangulation of D-{origin}. Since the triangulation was invariant it follows that U is invariant and $T \times T \mid U$ has period r and no fixed points. That U is homeomorphic to E^3 follows from:

LEMMA. *If K is a contractible complex (possibly infinite) having codimension at least 3 in a combinatorial n-manifold M, then the open star neighborhood U of K in the second barycentric subdivision of M is homeomorphic to E^n.*

Proof of the Lemma. Every closed simplex in $\mathrm{Cl}\, U$ is the joint of a subsimplex in K with a subsimplex in $\mathrm{Bd}\, U$, hence every point in $U - K$ determines a unique line segment $[x, y]$ in some simplex in $\mathrm{Cl}\, U$ having one endpoint in K and the other in $\mathrm{Bd}\, U$. Parametrize linearly all such line segments in a continuous manner over U by t values in $[0,1]$ so that $t = 0$ corresponds to the end-point in K. Define $\pi\colon U \to K$ so that each of the half-open line segments is taken onto its end-point in K.

Let L be any subset of K and s be in $(0,1)$. Define $L(s)$ to be the set of all those points in $\pi^{-1}(L)$ whose t value is no greater than s. Then if L is compact, $L(s)$ is also. Next define a homeomorphism $g_s\colon U \to U$, fixed on K, as follows. Consider a segment as before, parametrized by $[0,1]$. g_s maps this segment piecewise linearly by taking $[0, s]$ linearly onto $[0, 1-s]$ and $[s, 1]$ linearly onto $[1-s, 1]$.

Express K as the union of a sequence of finite subcomplexes $\{L_i\}$. Since K is a deformation retract of U it follows that U is contractible, and since U is open in M it is a combinatorial manifold, hence we can use the corollary in [4] to get a combinatorial n-cell C_1 in U containing L_1 in its interior.

Let L_2' be a finite subcomplex containing $L_2 \cup \pi(C_1)$. Use the corollary in [4] again to get an n-cell C_2' in U which contains L_2' in its interior. Since C_1 is contained in $L_2'(1-s)$ and $L_2'(s)$ is contained in the interior of C_2' for small enough s, it follows that for some s $g_s(C_2')$ will be an n-cell containing C_1 and $L_2(2/3)$ in its interior. Denote this n-cell by C_2.

Now proceed inductively getting n-cells $\{C_i\}$ in U such that $C_i \subset \mathrm{Int}\, C_{i+1}$ and $L_i(i/i+1) \subset C_i$. Then $U = \bigcup_{i=1}^{\infty} L_i(i/i+1) = \bigcup_{i=1}^{\infty} C_i$ and the last term is E^n by [1]. This proves the Lemma.

Finally, note that as an open set U inherits a differential structure from E^3 and relative to this structure $T \times T$ is differentiable, since $T \times T$ is orthogonal on E^3. Also U can be given an invariant triangulation in the following way. Consider $\mathrm{Cl}\, U$ as a subcomplex Q of the second barycentric subdivision of D-{origin}. Denote by U_1 the union of all closed simplexes in the barycentric subdivision Q_1 of Q which are contained in U. Denote by

Q_2 the subdivision of Q_1 obtained by starring all simplexes in Q_1 not in U_1. Let U_2 be the union of all simplexes in Q_2 contained in U. Continue on in this way getting $U = \bigcup_{i=1}^{\infty} U_i$ with each U_i invariant, and the collection of all simplexes occurring in the U_i give an invariant trangulation.

That these specific differential and combinatorial structures are equivalent to the standard ones follows from [7]. This finishes the proof of the theorm for $n = 8$. For $E^8 \times E^k$, define the action to be trivial on the second factor.

For each odd p let $T_{2p}\colon S^2 \to S^2$ be defined by $T_{2p} = T' \circ T_{2p}''$ where $T'\colon S^0 \to S^0$ interchanges points and $T_{2p}''\colon S^1 \to S^1$ rotates through $360/p$ degrees. Then T_{2p} has period $2p$ and no fixed point.

COROLLARY. *If $f\colon S^2 \to S^2$ is a map such that $T_{2p}f = fT_{2p}$, for some p, then f is essential.*

Proof. If there is a map f and an odd p such that f has degree 0 and $T_{2p}f = fT_{2p}$ then we assume (by taking the appropriate power of T) that p is prime, and the construction in the theorem can be redone using S^2 in place of S^8. This results in a differentiable map of period $2p$ on E^6 having no fixed points. This contradicts the previously mentioned result of Smith [6].

Conjecture. It seems plausible that $T_{2p}f = fT_{p2}$ implies degree of $f \equiv \pm 1 \pmod p$.

UNIVERSITY OF MICHIGAN AND
INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] M. Brown, "The monotone union of open n-cells is an open n-cell," *Proceedings of the American Mathematical Society*, vol. 12 (1961), pp. 812-814.

[2] P. E. Conner and E. E. Floyd, "On the construction of periodic maps without fixed points," *Proceedings of the American Mathematical Society*, vol. 10 (1959), pp. 354-360.

[3] J. M. Kister, "Examples of periodic maps on Euclidean spaces without fixed points," *Bulletin of the American Mathematical Society*, vol. 67 (1961), pp. 471-474.

[4] D. R. McMillan and E. C. Zeeman, "On contractible open manifolds," *Proceedings of the Cambridge Philosophical Society*, vol. 58 (1962), pp. 221-224.

[5] P. A. Smith, "Fixed point theorems for periodic transformations," *American Journal of Mathematics*, vol. 63 (1941), pp. 1-8.

[6] ———, "New results and old problems in finite transformation groups," *Bulletin of the American Mathematical Society*, vol. 66 (1960), pp. 401-416.

[7] J. Stallings, "The piecewise-linear structure of Euclidean space," *Proceedings of the Cambridge Philosophical Society*, vol. 58 (1962), pp. 481-488.

A REFINEMENT OF VAN DER CORPUT'S THEOREM ON CONVEX BODIES.*

By G. K. White.

1. **Introduction.** Let K be an open convex body in n-space, symmetric about O and with volume $V(K)$. Let Λ be any lattice with determinant $\Delta \neq 0$. Then according to an assertion of Blichfeldt [1], subsequently verified by Van der Corput [2], if $V(K) > 2^n k\Delta$ $(k = 1, 2, \cdots)$ then K contains at least k pairs $\pm P \neq O$ of points of Λ. I observe that the ideas in his proof give, in fact, a little more information about the points in K:

Theorem A. *If $V(K) > k2^n\Delta$, then there are k distinct points $P_i \neq O$ $(i = 1, \cdots, k)$ of Λ such that*

$$P_i \in K \quad (i = 1, \cdots, k), \text{ and} \tag{1}$$

$$P_i - P_j \in K \quad (i, j = 1, \cdots, k). \tag{2}$$

The following immediate consequences of Theorem A (for the case $k = 2$) may be of interest in application:

(i) If the points of Λ other than O in K form a linearly independent set of $t \leqq n$ pairs of points, then $V(K) \leqq 2^{n+1}\Delta$.

(ii) If all the points of Λ except O in K lie on two distinct hyperplanes $\sum a_i x_i = \pm a \neq 0$, then $V(K) \leqq 2^{n+1}\Delta$.

It may be remarked that the theorem can also be proved directly by a method of Mordell [3] which van der Corput [2] utilized to obtain his results. We offer a third proof, which has the advantage of giving additional information about the critical cases, and for this, it is convenient to have the main assertion in its negative form:

Theorem B. *If there are at most $k = 1, 2, \cdots$ distinct points $\{O = P_0, P_1, \cdots, P_{k-1}\}$ including O of Λ such that*

$$P_i \in K \quad (i = 0, 1, \cdots, k-1), \text{ and} \tag{3}$$

$$P_i - P_j \in K \quad (i, j = 0, 1, \cdots, k-1) \tag{4}$$

* Received January 11, 1963.

then

$$V(K) \leqq k2^n\Delta \tag{5}$$

and $V(K) = k2^n\Delta$ if and only if every grid of the form $\Gamma = X + \Lambda$ (where X is an arbitrary point in space) which has no point on the boundary of K, has exactly $k2^n$ points in K.

Before proceeding to the proof, we consider some properties of lattice-coverings of space which will be useful in characterization of the critical cases in Theorem B.

2. Lattice-coverings of bounded sets. In the course of our proof of Theorem B we shall use our main result (Theorem 1), which is applicable not only to K but also to more general bounded sets, and in preparation we introduce some symbols and definitions:

Let S be any bounded set of content $V = V(S) > 0$, and for $P \in \Lambda$ let $S(O) = S$, $S(P) = P + S(O)$ denote the respective translated sets. We thus induce a multiple *S-covering*[1]

$$\sum_{P \in \Lambda} S(P)$$

of space, and for any point X we set

$$k(X) = \text{Cardinal}\,\{P : X \in S(P)\}.$$

Since S is bounded, $k(X)$ is bounded and integrable. Let $M = M(S, \Lambda)$ denote the maximum value of $k(X)$. Then

$$0 \leqq k(X) \leqq M(S, \Lambda) = M \tag{6}$$

for all X in space, and we may define

$$\rho = \rho(S, \Lambda) = 1/V \int_S k(X)\,dX, \tag{7}$$

ρ representing the average density of the covering, taken over any one of the sets $S(P)$. Clearly,

$$1 \leqq \rho \leqq M. \tag{8}$$

THEOREM 1. *Let $M \geqq 1$ be any given integer and let Λ be a fixed lattice. Suppose that S is a bounded set in space with content $V(S) > 0$ and with $M(S, \Lambda) = M$. Let Z denote the set of all such S with fixed $\rho(S, \Lambda) = \rho$, say. Then*

---

[1] As distinct from an ordinary S-covering $\bigcup_{P \in \Lambda} S(P)$.

(9) $\sup_{S \in Z}{}^{2} V(S) = ([\rho]+1)[\rho]\Delta/(2[\rho]+1-\rho) \begin{cases} = \rho\Delta \textit{ if } \rho \textit{ is integral} \\ < \rho\Delta \textit{ otherwise,} \end{cases}$

where $\rho \leqq M$. *Also, for any* S *in* Z,

$$V(S) = \rho\Delta \text{ if and only if } k(X) = \rho P\cdot P.^{3} \tag{10}$$

COROLLARY 1. *Suppose that* S *is open and bounded, and that the boundary of* S *has content* 0. *Then* (10) *may be replaced by*

$V(S) = \rho\Delta$ *if and only if every grid of the form* $\Gamma = X + \Lambda$ (*where* X *is an arbitrary point in space*) *which has no point on the boundary of* S, *has exactly* ρ *points in* S.

COROLLARY 2. *Suppose* $O \in S$ *and that* S *is open, symmetric about the origin and bounded, and that the boundary of* S *has content* 0, *and let* Γ *be any one of the* 2^n *grids of the form* $\Gamma = B + \Lambda$, *with* $2B \in \Lambda$. *Then in each limiting case* $V(S) = \rho\Delta$, *there are points of* Γ *on the boundary of* S, *when*

$$\Gamma = \Lambda, \textit{ and } \rho \textit{ is even}$$

or when

$$\Gamma \neq \Lambda, \textit{ and } \rho \textit{ is odd.}$$

We note that since $k(X)$ is integral, the limiting cases (10) can occur only when ρ is an integer.

I should like to thank Dr. J. H. H. Chalk here for his valuable assistance throughout the preparation of this paper.

3. Proof of Theorem 1. Without loss of generality, Λ is the lattice of points with integral coordinates. Then, employing the usual notation, $\Lambda = \Lambda_0$. The coordinates of any point X of space shall be denoted by $(\xi_1, \cdots, \xi_n)$. For j any positive integer, we consider the half-open integral box

$$B_j = \{X : -j/2 \leqq \xi_i < j/2\} \qquad (i = 1, \cdots, n)$$

and define

$$S_i = \{X : k(X) = i\}$$

$$W_i = \int_{B_1 \cap S_i} dX \qquad (i = 0, 1, \cdots, M)$$

$$U_i = \int_{S \cap S_i} dX.$$

---

[2] The supremum is attained.

[3] $P \cdot P$ (almost everywhere) shall be taken in this paper to mean "everywhere in space, except on a set whose intersection with any bounded region has content 0."

We first show that

(11) $$U_i = iW_i \qquad (i = 0, 1, \cdots, M).$$

Since S is bounded, we may enclose it in some box B_j. If, for any $m > j + 1$, we form the sum

(12) $$\sum_{P \in B_m} S(P) \cap S_i,$$

points of S_i outside B_{m+j} will not appear in (12), while those inside will be counted at most i times, points in B_{m-j} being counted exactly i times. Since there are exactly $iN^n W_i$ points of S_i inside B_N $(N = 1, 2, \cdots)$, we have

$$i(m-j)^n W_i \leqq m^n U_i \leqq i(m+j)^n W_i.$$

If we now let $m \to \infty$, (10) follows. Next, we observe that clearly,

$$1 = \sum_i W_i \text{ and } V = \sum_i iW_i = \int_{B_1} k(X)\, dX.$$

Therefore the content V of S may be taken to represent the average density of the S-covering $\sum_{P \in \Delta_0} S(P)$ of space. By (7),

$$\rho V = \int_S k(X)\, dX - \sum_i \int_{S \cap S_i} k(X)\, dX = \sum_i iU_i - \sum_i i^2 W_i.$$

We thus have, with a convenient change of notation,

(13) $$W_i = x_i^2 > 0, \qquad (i = 0, 1, \cdots, M)$$

(14) $$V = \sum_i ix_i^2,\ \rho V = \sum_i i^2 x_i^2,\ 1 = \sum_i x_i^2 \qquad (i = 0, 1, \cdots, M).$$

We now find the maximum V for fixed ρ, under conditions (14). First, suppose $x_0 \neq 0$. Then we may replace x_i^2 by $y_i^2 = x_i^2(1+\epsilon)$ $(i = 1, \cdots, M)$ and x_0^2 by $y_0^2 = 1 - \sum_{i=1}^{M} x_i^2(1+\epsilon)$, for some $\epsilon > 0$, and thus obtain a larger V. So without loss of generality we set $x_0 = 0$, and rewrite (14)

(15) $$\left.\begin{aligned} V &= \sum ix_i^2 \\ 0 &= \sum (i^2 - \rho i) x_i^2 \\ 1 &= \sum x_i^2 \end{aligned}\right\} \qquad (i = 1, \cdots, M),$$

and for $M > 2$ we define

$$E(\lambda, \mu) = \sum_{i=1}^{M} ix_i^2 + \lambda \sum_{i=1}^{M} (i^2 - \rho i) x_i^2 + \mu \sum_{i=1}^{M} x_i^2.$$

For the maximum of V under the restrictions $0=\sum(i^2-\rho_i)x_i^2$, $1=\sum x_i^2$, we have by the method of Lagrange multipliers

$$0=\partial E/\partial x_i=2(i+\lambda(i^2-\rho i)+\mu)x_i^2. \tag{16}$$

Since ρ is fixed, it is easy to see that the quadratic in i, $i+\lambda(i^2-\rho i)+\mu=0$, cannot be satisfied by three distinct values of i. Thus there exist integers r, s $(1\leqq r<s\leqq M)$ for which $x_i=0$ unless $i=r$ or s, and (15) becomes

$$\begin{aligned} V&=rx_r^2+sx_s^2\\ 0&=(r^2-\rho r)x_r^2+(s^2-\rho s)x_s^2 \qquad (1\leqq r<s\leqq M)\\ 1&=x_r^2+x_s^2. \end{aligned} \tag{17}$$

We distinguish two cases:

First, suppose that $x_r\neq 0$, $x_s\neq 0$. Then since

$$0=(r-\rho)rx_r^2+(s-\rho)sx_s^2$$

and $1\leqq r<s$, we have

$$r<\rho<s, \tag{18}$$

and so $r+s-\rho>0$. Solving (17), we find

$$\begin{aligned} V&=rs/(r+s-\rho) \qquad (19)\\ &=\rho+(r-\rho)(s-\rho)/(r+s-\rho)\\ &<\rho, \text{ by } (18). \qquad (20) \end{aligned}$$

We rewrite $(r-\rho)(s-\rho)/(r+s-\rho)=-(\rho-r)/(1+r(s-\rho)^{-1})$. If we keep r fixed and let s vary $(\rho<s\leqq M)$, it is evident from (18), (19) that V will increase as s decreases, attaining its maximum at $s=[\rho]+1$. By a similar argument the maximum V for fixed s is at $r=[\rho]$. Thus the maximum V is

$$V=([\rho]+1)[\rho]/(2[\rho]+1-\rho) \tag{21}$$

under our first assumption that $x_r\neq 0$, $x_s\neq 0$.

For our second case, we suppose the contrary. Then without loss of generality $x_s=0$, and by (17), $1=x_r^2$ and

$$V=r=\rho. \tag{22}$$

Thus this case arises if and only if the sets $\{S(P): P\in\Lambda_0\}$ cover space exactly $r=\rho$ times, almost everywhere, and so (10) follows from (20), (22). By

(22), this occurs only for ρ an integer (r), so for non-integral ρ (21) gives the overall maximum of V, which is less than ρ by (20). For integral ρ, the expression in (21) is equal to ρ, and hence by (22) is again the overall maximum. We have thus verified (9), and our proof is complete.

Proof of Corollary 1. Clearly,

$$X \in S(P) \text{ if and only if } X - P \in S(O) = S.$$

So if we define $N(\Gamma)$ to be the number of points in S of the grid $\Gamma = X + \Lambda_0$, it is immediate that

$$N(\Gamma) = k(X). \tag{23}$$

Assume now that $k(X) = \rho P \cdot P$, and Γ does not intersect the boundary of any $S(P)$; we deduce that $N(\Gamma) = \rho$. For $X \in \Gamma$, X will be in exactly $k(X)$ of the bodies $S(P)$, but not on the boundary of any. Since S is open, there clearly exists some neighbourhood $\mathfrak{N}(X)$ of X with the same property. $\mathfrak{N}(X)$ has non-zero content, so if $V(S) = \rho\Delta$, we have by Theorem 1 that for every point Y in $\mathfrak{N}(X)$, $k(Y) = \rho$. In particular, with $Y = X$ we have $N(\Gamma) = k(X) = \rho$, by (23).

Conversely, assume that for some positive integer k, $N(\Gamma) = k$ for every grid Γ which does not intersect the boundary of any $S(P)$; we deduce $k(X) = \rho P \cdot P$. Let T be the set of all points X whose corresponding grid $\Gamma = X + \Lambda_0$ has the above property, and let T' be its complement. If X is in the unit box B_1, clearly $X \in T'$ if and only if X lies on the boundary of some $S(P)$. Only a finite number of the bounded sets $S(P)$ can intersect B_1, and the boundary of any $S(P)$ has content 0, by the hypothesis of the corollary. Hence it follows that $B_1 \cap T'$ has content 0, and $k(X) = kP \cdot P$. It remains to show that $k = \rho$. We have

$$\begin{aligned} \rho V(S) &= \int_{S \cap T} k(X)\,dX + \int_{S \cap T'} k(X)\,dX \\ &= kV(S \cap T) + \alpha V(S \cap T') \quad (0 \leqq \alpha \leqq M \text{ by } (6)) \\ &= kV(S) + 0, \text{ since } V(B_1 \cap T') = 0. \end{aligned}$$

Thus $k = \rho$, and a reference to Theorem 1 (10) completes our proof.

Proof of Corollary 2. Let Γ be of the form $\Gamma = B + \Lambda$, where $2B \in \Lambda$. Clearly $X \in \Gamma$ implies $2X \in \Lambda$, and thus $-X = X + (-2X) \in \Gamma$. Also $X \in S$ implies $-X \in S$, since S is symmetric about the origin. Thus all points $X \neq O$ of $\Gamma \cap S$ occur in pairs $(X, -X)$, and so since $O \in S$, the number of points of Γ must be odd for $\Gamma = \Lambda$, even otherwise. If we now assume we have a

limiting case $V(S) = \rho\Delta$, and that there are no points of Γ on the boundary of S, then there must be exactly ρ points of Γ inside S, by Corollary 1. Our result follows easily, by contradiction.

4. Proof of Theorem B. Suppose $V(K) > k2^n\Delta$, and let $K' = \frac{1}{2}K$, so that $V(K') > k\Delta$. Then by Theorem 1, with $S = K'$,

$$k\Delta < V(K') \leqq \rho\Delta \leqq \sup_X k(X)\Delta,$$

so there must be some point X contained in at least $k+1$ bodies

$$K'(Q_i)\,(Q_i \in \Lambda, i = 0, 1, \cdots, k).$$

Thus

$$K'(Q_0) \cap K'(Q_1) \cap \cdots \cap K'(Q_k) \neq \emptyset,$$

where $Q_0, Q_1, \cdots, Q_k$ are distinct points of Λ. Clearly, from the symmetry and convexity of K,

$$P_i = Q_i - Q_0 \in K(O) \qquad (i = 0, 1, \cdots, k)$$
$$P_i - P_j \in K(O) \qquad (i, j = 0, 1, \cdots, k)$$

contrary to our original hypothesis. Thus,

$$V(K) \leqq k2^n\Delta.$$

The boundary of the convex body K is a single Jordan arc, and K is open. The remainder of Theorem B is thus a direct consequence of Corollary 1.

UNIVERSITY OF TORONTO.

REFERENCES.

[1] H. F. Blichfeldt, "Notes on geometry of numbers," *Bulletin of the American Mathematical Society*, vol. 27 (1921), pp. 150, 152-3.

[2] J. G. van der Corput, "Verallgemeinerung einer Mordellschen Beweismethode in der Geometrie der Zahlen," *Acta Arithmetica*, vols. 1, 2 (1935, 1936), pp. 62-66, 145-146.

[3] L. J. Mordell, "On some arithmetical results in the geometry of numbers," *Compositio Mathematica*, vol. 1 (1934), pp. 248-253.

CALCUTTA

LIE GROUP REPRESENTATIONS ON POLYNOMIAL RINGS.*[1]

By BERTRAM KOSTANT.[2]

0. Introduction. 1. Let G be a group of linear transformations on a finite dimensional real or complex vector space X. Assume X is completely reducible as a G-module. Let S be the ring of all complex-valued polynomials on X, regarded as a G-module in the obvious way, and let $J \subseteq S$ be the subring of all G-invariant polynomials on X.

Now let J^+ be the set of all $f \in J$ having zero constant term and let $H \subseteq S$ be any graded subspace such that $S = J^+S + H$ is a G-module direct sum. It is then easy to see that

$$S = JH. \tag{0.1.1}$$

(Under mild assumptions H may be taken to be the set of all G-harmonic polynomials on X. That is, the set of all $f \in S$ such that $\partial f = 0$ for every homogeneous differential operator ∂ with constant coefficients, of positive degree, that commutes with G.)

One of our main concerns here is the structure of S as a G-module. Regard S as a J-module with respect to multiplication. Matters would be considerably simplified if S were free as a J-module. One shows easily that S is J-free if and only if $S = J \otimes H$. This, however, is not always the case. For example S is not J-free if G is the two element group $\{I, -I\}$ and $\dim X \geqq 2$. On the other hand one has

Example 1. It is due to Chevalley (see [2]) that if G is a finite group generated by reflections then indeed $S = J \otimes H$. Furthermore the action of G on H is equivalent to the regular representation of G.

Example 2. S is J-free in case G is the full rotation group (with respect to some Euclidean metric on X. For convenience assume in this example that $\dim X \geqq 3$). Note that the decomposition of a polynomial according to the relation $S = J \otimes H$ is just the so-called "separation of variables" theorem for polynomials. This is so because J is the ring of radical polynomials and H is the space of all harmonic polynomials (in the usual sense).

* Received February 11, 1963.

[1] This research was supported in part by National Science Foundation Grant NSF-G19992.

[2] The author is an Alfred P. Sloan Fellow.

Now, for any $x \in X$, let $O_x \subseteq X$ denote the G-orbit of x and let $S(O_x)$ be the ring of all functions on O_x defined by restricting S to O_x. Since J reduces to constants on any orbit it follows that (0.1.1) induces a G-module epimorphism

(0.1.2) $$H \to S(O_x).$$

Since our major concern is the case where X is a reductive Lie algebra and G is the adjoint group and since the methods used there belong to algebraic geometry we will assume now that X is complex and the G is algebraic and reductive. All varieties considered are over $\boldsymbol{C}$. If Y has an algebraic structure $R(Y)$ will denote the ring of everywhere defined rational functions in Y. Obviously one always has

(0.1.3) $$S(O_x) \subseteq R(O_x).$$

On the other hand if $G^x \subseteq G$ is the isotropy group defined by $x \in X$ then one has a G-module isomorphism

(0.1.4) $$R(G/G^x) \to R(O_x).$$

The significance of (0.1.4) is that one knows the G-module structure of $R(G/G^x)$ completely by a very simple algebraic Frobenius reciprocity theorem (even though G^x may not be reductive). In fact if V^λ is any irreducible G-module with respect to the representation ν^λ and V_λ is the dual module then one has

(0.1.5) $$\text{mult. of } \nu^\lambda \text{ in } R(G/G^x) = \dim V_\lambda{}^{G^x}$$

where $V_\lambda{}^{G^x}$ is the space of vectors in V_λ fixed under G^x.

Now in Examples 1 and 2 (assume complexified) the following three optimum situations occur:

(a) S is J-free so that $S = J \otimes H$

(b) the map $H \to S(O_x)$ is an isomorphism for certain $x \in X$ and for those x

(c) $R(O_x) = S(O_x)$.

But one observes that if in any general case, (b) and (c) hold then, clearly, upon combining (0.1.4) and (0.1.5) one gets the G-module structure of H. If one gets in addition the "graded" G-module structure of H and knows the structure of J then one gets the full graded G-module structure of S in case (a) also holds.

In Example 2 the conditions (b) and (c) hold for any $x \neq 0$ (even if $(x, x) = 0$). In fact, classically, one has exploited (b) and (c) for $(x, x) > 0$

to solve the Dirichlet problem with the sphere as boundary. That is, if f is any continuous function on the sphere one first expands f as a Fourier development of spherical harmonics f_m. The sphere is $O_x \cap \boldsymbol{R}^n$ and the f_m are in $R(O_x)$. The equality $R(O_x) = S(O_x)$ and the isomorphism $H \to S(O_x)$ then yields the extension of f_m uniquely as harmonic polynomials h_m on X. But this yields the desired extension of f.

In Example 1 the conditions (b) and (c) are satisfied for any "regular" element $x \in X$.

Our first concern in this paper is to give criteria for (a), (b) and (c) to hold in general. Since our interest is in the continuous case we will assume G is connected (and hence a variety). Thus Example 2 rather than Example 1 serves as a model.

Now let $P \subseteq X$ be the cone of common zeros defined by the ideal J^+S in S. Let X^* be the dual space to X and let $P^* \subseteq X^*$ be defined in a similar way with the roles of X and X^* interchanged. As a criterion to establish (a) and more we prove

PROPOSITION 0.1. *Assume* (1) *that J^+S is a prime ideal in S and* (2) *there exists an orbit $O_e \subseteq P$ which is dense in P. Then $S = J \otimes H$. Furthermore if G is a subgroup of the complex rotation group then H may be taken as the space of all G-harmonic polynomials. Moreover H then coincides with the space spanned by all powers f^k where $f \in P^*$.*

It may be observed that the criterion is satisfied in Example 2.

An element $x \in X$ is called quasi-regular if $P \subseteq \overline{\boldsymbol{C}^* \cdot O_x}$. A criterion to establish (b) is given by

PROPOSITION 0.2. *Assume conditions* (1) *and* (2) *of Proposition* 0.1 *are satisfied. Then the G-module epimorphism $H \to S(O_x)$ is an isomorphism for any quasi-regular element $x \in X$.*

It may be observed that in Example 2 every nonzero $x \in X$ is quasi-regular.

From known facts in algebraic geometry one has the following criterion to insure (c).

PROPOSITION 0.3. *Let $x \in X$ and assume* (1) *the closure $\bar{O}_x$ is a normal variety and* (2) *$\bar{O}_x - O_x$ has a codimension of at least* 2 *in $\bar{O}_x$. Then $R(O_x) = S(O_x)$.*

It may be observed that the conditions of Proposition 0.3 are satisfied for every $x \in X$ in Example 2.

Now assume that $X = \mathfrak{g}$ is a complex reductive Lie algebra and G is the adjoint group. Here the structure of J is given by a theorem of Chevalley. This aserts that J is a polynomial ring in l (the rank of $\mathfrak{g}$) homogeneous generators u_i, $i = 1, 2, \cdots, l$ with $\deg u_i = m_i + 1$ where the m_i are the exponents of $\mathfrak{g}$.

Now one knows that here P is the set of all nilpotent elements of $\mathfrak{g}$ ([13], Theorem 9.1). But then by [13], Corollary 5.5, P does contain a dense orbit O_e, namely, the set of all principal nilpotent elements in $\mathfrak{g}$. Thus to apply Propositions 0.1 and 0.2 one must prove that J^+S is a prime ideal.

If $n = \dim \mathfrak{g}$ (all dimensions are over $\boldsymbol{C}$) then one sees easily that $n - l$ is the maximal dimension of any orbit. Let $\mathfrak{r} = \{x \in \mathfrak{g} \mid \dim O_x = n - l\}$. Any regular element $x \in \mathfrak{g}$ belongs to $\mathfrak{r}$. But also $e \in \mathfrak{r}$ for any principal nilpotent element. These in fact are extreme cases.

PROPOSITION 0.4. *Let $x \in \mathfrak{g}$ be arbitrary. Write (uniquely) $x = y + z$ where y is semi-simple, z is nilpotent and $[y, z] = 0$. Let $\mathfrak{g}^y$ be the centralizer of y in $\mathfrak{g}$ so that $\mathfrak{g}^y$ is a reductive Lie algebra and $z \in \mathfrak{g}^y$. Then $x \in \mathfrak{r}$ if and only if z is principal nilpotent in $\mathfrak{g}^y$.*

Let $x \in \mathfrak{g}$. Consider the values $(du_i)_x$ of the l-differential forms du_i, $i = 1, 2, \cdots, l$, at x. It is known that these covectors are linearly independent whenever x is regular. (One recalls that the product of the positive roots is the determinant of an $l \times l$ minor of a certain $n \times l$ matrix determined by the du_i.) But to prove the primeness of the ideal J^+S one needs to know that these covectors are linearly independent if x is a principal nilpotent element. This fact is contained in

THEOREM 0.1. *Let $x \in \mathfrak{g}$. Then the $(du_i)_x$ is linearly independent if and only if $x \in \mathfrak{r}$.*

Proposition 0.1 may now be applied.

THEOREM 0.2. *One has $S = J \otimes H$ where H is the space of all G-harmonic polynomials on $\mathfrak{g}$. Furthermore H coincides with the space of all polynomials spanned by all powers of "nilpotent" linear functionals.*

Since Theorem 0.1 shows also that P is a complete intersection the decomposition $S = J \otimes H$ when combined with Proposition 5, § 78, in FAC [15], gives, in the notation of FAC, all the sheaf cohomology groups $H^j(\boldsymbol{P}, \boldsymbol{\mathcal{O}}(m))$ where $\boldsymbol{P}$ is the projective variety defined by P.

Another application of the primeness of J^+S in algebraic geometry is

THEOREM 0.3. *The intersection multiplicity of P, at the origin, with any Cartan subalgebra is w, where w is the order of the Weyl group.*

Next, Proposition 0.2 is put into effect for all orbits of maximal dimension by

THEOREM 0.4. *The set* $\mathfrak{r}$ *coincides with the set of all quasi-regular elements in* $\mathfrak{g}$. *(Thus* H *and* $S(O_x)$ *are isomorphic as G-modules for any* $x \in \mathfrak{r}$.)

As a consequence of Theorems 0.2 and 0.4 one shows that not only is the ideal J^+S prime in S but J_1S is prime for any prime ideal $J_1 \subseteq J$. Furthermore one gets the following characterization of all the invariant prime ideals in S which are generated by elements of J.

THEOREM 0.5. *Let* $I \subseteq S$ *be any G-invariant prime ideal. Let* $\mathfrak{u} \subseteq \mathfrak{g}$ *be the affine variety of zeros of* I. *Then* I *is of the form* $I = J_1S$, *for* J_1 *a prime ideal in* J, *if and only if* $\mathfrak{u} \cap \mathfrak{r}$ *is not empty.*

Since $R(O_x) = S(O_x)$ in case O_x is closed and since O_x is closed if x is regular one gets the G-module structure of H by applying Theorem 0.4 and (0.1.5) for x regular. Thus if D denotes the set of dominant integral forms corresponding to a Cartan subgroup A, so that D indexes all the irreducible representations of G as highest weights, then one has

$$\text{(0.1.6)} \qquad \text{mult } \nu^\lambda \text{ in } H = l_\lambda$$

where $l_\lambda = \dim V_\lambda{}^A$ is the multiplicity of the zero weight of ν_λ.

In order to determine the G-module structure of S^k, the space of homogeneous polynomials on $\mathfrak{g}$ of degree k, one must know more than (0.1.6). In fact using the relation $S = J \otimes H$ what one wants is the multiplicity of ν^λ in $H^j = S^j \cap H$ for any λ and j. As it turns out, for this, one needs $R(O_e) = S(O_e)$ where e is a principal nilpotent element. To show the latter using Proposition 0.3 it is enough to show that P is a normal variety and $P - O_e$ has a codimension of at least 2 in P.

Let $\boldsymbol{\mathcal{O}}_\mathfrak{r}$ be the set of all orbits of maximal dimension $(n - l)$. The set $\boldsymbol{\mathcal{O}}_\mathfrak{r}$ may be parametrized by $\boldsymbol{C}^l$ in the following way. Let

$$\boldsymbol{u} \colon \mathfrak{g} \to \boldsymbol{C}^l$$

be the morphism given by putting $u(x) = (u_1(x), \cdots, u_l(x))$ for any $x \in \mathfrak{g}$. Since $\boldsymbol{u}$ reduces to a constant on any orbit it induces a map

$$\eta_\mathfrak{r} \colon \boldsymbol{\mathcal{O}}_\mathfrak{r} \to \boldsymbol{C}^l.$$

It is known that u induces a bijection from the set of all orbits consisting of semi-simple elements onto $\boldsymbol{C}^l$ (for completeness a proof of this fact will be given here). Combining this with Proposition 0.4 one obtains

THEOREM 0.6. *$\eta_{\mathfrak{r}}$ is a bijection.*

Thus to each $\xi \in \boldsymbol{C}^l$ there exists a unique orbit, $O^{\mathfrak{r}}(\xi)$, of dimension $n-l$ which correspond to ξ under $\eta_{\mathfrak{r}}$. Now let $P(\xi)=u^{-1}(\xi)$ for any $\xi \in \boldsymbol{C}^l$ so that

$$\mathfrak{g}=\bigcup_{\xi \in C^l} P(\xi)$$

is a disjoint union. Note that $P(\xi)=P$ and $O^{\mathfrak{r}}(\xi)=O_e$ if ξ is the origin of $\boldsymbol{C}^l$. One proves

THEOREM 0.7. *For any $\xi \in \boldsymbol{C}^l$ one has*

$$P(\xi)=\overline{O^{\mathfrak{r}}(\xi)}$$

so that $P(\xi)$ is a variety of dimension $n-l$. Moreover $P(\xi)$ is a complete intersection and $O^{\mathfrak{r}}(\xi)$ coincides with the set of non-singular points on $P(\xi)$. Finally $P(\xi)$ is a finite union of orbits so that $\bar{O}_x$ is a finite union of orbits for any $x \in \mathfrak{g}$.

Since $P(\xi)$ is a complete intersection and since its singular locus is the complement (a finite union of orbits) of $O^{\mathfrak{r}}(\xi)$ in $P(\xi)$ one would get the normality of $P(\xi)$ by a theorem of Seidenberg if one knew the dimension of the other orbits in $P(\xi)$ were at most $n-l-2$.

Now it is well known that $\dim O_x$ is even (and hence $\dim_{\boldsymbol{R}} O_x$ is a multiple of 4) for any semi-simple element $x \in \mathfrak{g}$. Less known is the following proposition observed independently by the author, Borel, and (most simply proved by) Kirillov.

PROPOSITION 0.5. *The dimension of O_x is even for any $x \in \mathfrak{g}$.*

Combining Theorem 0.7 and Proposition 0.5 one obtains

THEOREM 0.8. *Let $\xi \in \boldsymbol{C}^l$ be arbitrary. Then $P(\xi)$ is a normal variety and the codimension of $P(\xi)-O^{\mathfrak{r}}(\xi)$ in $P(\xi)$ is at least 2.*

Applying Proposition 0.3 one then has

THEOREM 0.9. *Let $x \in \mathfrak{r}$. Then $R(O_x)=S(O_x)$. (This implies that all $R(O_x)$ for $x \in \mathfrak{r}$ are isomorphic as G-modules; even though they are not in general as rings.) Let $\xi=u(x)$. Then $R(O_x)$ $(=R(G/G^x))$ is an affine algebra (even though O_x is not necessarily an affine variety) and $P(\xi)$ is the variety of all maximal ideals of $R(O_x)$. Thus the embedding of G/G^x in $\mathfrak{g}$ as O_x is special in that any morphism of G/G^x (or O_x) into any affine variety extends uniquely to a morphism of $P(\xi)=\bar{O}_x$ into the variety. (In particular*

this holds for O_e *and* $\bar{O}_x = P$.) *Finally* (*using* (0.1.5) *and the equality* $R(O_x) = S(O_x)$) *one has, for any* $\lambda \in D$

$$\dim V_\lambda{}^{G^x} = l_\lambda \tag{0.1.7}$$

so that the left side of (0.1.7) *is independent of* $x \in \mathfrak{r}$.

Now let $\{e_-, x_0, e\}$ be a principal S-triple (that is, a "canonical" basis of a principal three dimensional simple Lie subalgebra). In particular then e is a principal nilpotent element. Used heavily in the theorems above is the result of [13] which asserts that $\mathfrak{g}^e$ is l-dimensional and has a basis z_i, $i = 1, 2, \cdots, l$, such that

$$[x_0, z_i] = m_i z_i \tag{0.1.8}$$

where, we recall, the m_i are the exponents of $\mathfrak{g}$. The main application of this is the following result: Let $\mathfrak{a}$ be any subspace of $\mathfrak{g}$ such that (1) $\mathfrak{g} = \mathfrak{a} + [e_-, \mathfrak{g}]$ is a direct sum and (2) $\mathfrak{a}$ is stable under $\operatorname{ad} x_0$ (e.g. take $\mathfrak{a} = \mathfrak{g}^e$). Then if $\mathfrak{v}$ is the l-plane defined by the translation $\mathfrak{v} = e_- + \mathfrak{a}$ one has

THEOREM 0.10. *The variety* $\mathfrak{v}$ *is contained in* $\mathfrak{r}$. *Moreover each orbit in* $\mathcal{O}_\mathfrak{r}$ *intersects* $\mathfrak{v}$ *in one and only one point. Finally the mapping* $f \to f \mid \mathfrak{v}$ *induces an isomorphism of* J *onto* $R(\mathfrak{v})$.

Remark. If $\mathfrak{g}$ is the set of all $l \times l$ complex matrices then one shows easily that $\mathfrak{r}$ is the set of all matrices whose characteristic polynomial is equal to their minimal polynomial. An example of the subvariety $\mathfrak{v}$ is the set of all "companion" matrices. Here the validity of Theorem 0.10 is a well-known fact in matrix theory.

Now since $\mathfrak{g}^e = \mathfrak{g}^{G^e}$ (because $\mathfrak{g}^e$ is commutative) and since (0.1.7) holds for $x = e$ this suggests a generalization of the notion of exponent. Let V be any finite dimensional G-module with respect to a representation ν. If l^ν is the multiplicity of the zero weight of ν then by (0.1.7) one has $\dim V^{G^e} = l^\nu$. It follows therefore that there exists a unique non-decreasing sequence of non-negative integers $m_i(\nu)$, $i = 1, 2, \cdots, l^\nu$, such that one has

$$\nu(x_0) z_i = m_i(\nu) z_i$$

for a basis z_i of V^{G^e}. If ν is the adjoint representation the $m_i(\nu)$ are the usual exponents. If $\nu = \nu^\lambda$ we will write $m_i(\lambda)$ for $m_i(\nu^\lambda)$ and note (because the highest weight has multiplicity one) that

$$m_j(\lambda) = o(\lambda) \text{ for } j = l_\lambda$$

where $o(\lambda)$ is the sum of the coefficients of λ relative to the simple roots and

that this highest value occurs with multiplicity one among the generalized exponents $m_i(\lambda)$. (This specializes to the familiar relation $m_l = o(\psi)$ when $\mathfrak{g}$ is simple and ψ is the highest root.)

The following theorem now gives the G-module structure of H^j and hence S^k for any j and k.

THEOREM 0.11. *Let $\lambda \in D$ be arbitrary and let $H(\lambda)$ be the set of G-harmonic polynomials which transform under G according to ν^λ. Let (by (0.1.6)) $H(\lambda) = \sum_{j=1}^{l_\lambda} H_j(\lambda)$ be a decomposition into irreducible components so that $H_j(\lambda) \subseteq H^{n_j}$ where n_j, $j = 1, 2, \cdots, l_\lambda$, is a non-decreasing sequence of integers. Then $n_j = m_j(\lambda)$ for all j. In particular then $k = o(\lambda)$ is the highest degree k such that ν^λ occurs in H^k. Moreover it occurs with multiplicity one for this value of k.*

Assume for convenience that $\mathfrak{g}$ is simple and let $\psi \in D$ be the highest root. Let x_i, $i = 1, 2, \cdots, n$ be a basis of $\mathfrak{g}$. If the $u_j \in J$ are chosen properly one sees that $\dfrac{\partial u_j}{\partial x_i}$, $i = 1, 2, \cdots, n$, is a basis of $H_j(\psi)$. One notes then that Theorem 0.11 is a generalization of the result in [13] given by (0.1.8).

H. S. M. Coxeter observed and A. J. Coleman proved in [4] that if W is the Weyl group and $\sigma \in W$ is the Coxeter-Killing transformation then the eigenvalues of σ operating on the Cartan subalgebra are $e^{2\pi i m_j/s}$, $j = 1, 2, \cdots, l$, where s is order of σ. Now more generally W operates on the zero weight space of V^λ for any $\lambda \in D$ according (say) to some representation π^λ of W. As a generalization of the Coxeter-Coleman theorem one now has

THEOREM 0.12. *For any $\lambda \in D$ the eigenvalues of $\pi^\lambda(\sigma)$ are $e^{2\pi i m_j(\lambda)/s}$, $j = 1, 2, \cdots, l_\lambda$.*

0.2. By applying the Birkhoff-Witt theorem the results above carry over from S to U, the universal enveloping of $\mathfrak{g}$ (U is obviously a G-module in a natural way).

THEOREM 0.13. *Let U be the universal enveloping algebra over $\mathfrak{g}$ and let $Z \subseteq U$ be the center of U. Then U is free as a Z-module (under multiplication). In fact*

$$U = Z \otimes E \tag{0.2.1}$$

where E is the subspace (and G-submodule) of U spanned by all powers x^k for all nilpotent elements $x \in \mathfrak{g}$. Moreover E is equivalent to H as a G-module so that every irreducible representation of G occurs with finite multiplicity in E (in fact ν^λ occurs l_λ times in E for any $\lambda \in D$).

Let V be a finite dimensional irreducible U-module so that one has a G-module algebra epimorphism

$$\rho\colon U \to \operatorname{End} V$$

Since $\rho(Z)$ reduce to the scalars it follows from (0.2.1) that $\rho(E) = \operatorname{End} V$. Now let Y be any subspace of U. If Y is one-dimensional then it is due to Harish-Chandra that there exists an irreducible U-module V such that ρ is faithful on Y. This is not true in general if $\dim Y \geqq 2$. However it is true if $Y \subseteq E$.

THEOREM 0.14. *Let $Y \subseteq E$ be any finite dimensional subspace. Then there exists an irreducible U-module V such that ρ is faithful on Y.*

I would like to express my thanks to C. Chevalley, M. Rosenlicht and A. Seidenberg for helpful conversations about questions in algebraic geometry. In particular to Siedenberg for making me aware of his criterion for normality and to Chevalley for simplifying my proof of the primeness of J^+S.

1. Consequences of the primeness of J^+S and a dense orbit in P.
1. Let X be a n-dimensional vector space over the complex numbers C. Let $S_* = S_*(X)$ symmetric algebra over X. One knows that S_* may be regarded as the algebra of all differential operators ∂ on X which may be put in the form

$$\partial = \sum a_{i_1\cdots i_n}\left(\frac{\partial}{\partial z_1}\right)^{i_1}\cdots\left(\frac{\partial}{\partial z_n}\right)^{i_n}$$

where the $a_{i_1\cdots i_n}$ are complex constants and $z_1,\cdots,z_n$ are the affine coordinates of X.

Let $S^* = S^*(X)$ (or just S) be the symmetric algebra over the dual space to X. Then S is just the ring of all polynomials on X. In fact we take the point of view that X is an affine variety (over C) and S is its ring of everywhere defined rational functions. The algebra S (resp. S_*) is graded in the obvious way and a subspace $L \subseteq S$ (resp. $L \subseteq S_*$) will be called graded if it is spanned by its homogeneous components $L^j = L \cap S^j$ (resp. $L \cap S_j$).

Now one knows that a non-singular pairing of S_* and S into C is established by putting

$$\langle \partial, f\rangle = \partial f(0) \tag{1.1.1}$$

where $\partial \in S_*$, $f \in S$ and $\partial f(0)$ denotes the value of the function ∂f at the origin. In this way S_k is orthogonal to S^j if $j \neq k$ and becomes isomorphic to its dual if $k = j$. It is obvious from (1.1.1) that

$$\langle \partial_1\partial_2, f\rangle = \langle \partial_2, \partial_1 f\rangle \tag{1.1.2}$$

for any $\partial_1, \partial_2 \in S_*$ and $f \in S$ and hence in particular if $f \in S^m$ and $x \in X$ then by the Taylor expansion

$$\langle \frac{(\partial_x)^m}{m!}, f \rangle = f(x) \tag{1.1.3}$$

where ∂_x is the element of $S_1(X) \cong X$ corresponding to x.

Now assume that $G \subseteq \operatorname{Aut} X$ is a connected linear reductive algebraic group, i.e., G is the complexification of a connected compact subgroup of $\operatorname{Aut} X$. We regard G as not only operating on X, but by unique extension, as a group of algebra automorphisms of $S_*(X)$ and also as a group of algebra automorphisms of $S(X)$. The action on the latter is also uniquely defined by requiring that

$$\langle a \cdot \partial, a \cdot f \rangle = \langle \partial, f \rangle \tag{1.1.4}$$

for all $a \in G$, $\partial \in S_*$ and $f \in S$. Note by (1.1.3) that

$$(a \cdot f)(x) = f(a^{-1}x) \tag{1.1.5}$$

for any $x \in X$, $f \in S$ and $a \in G$.

Now let $J \subseteq S$ be the graded subring of G-invariant polynomials in S. That is

$$J = \{f \in S \mid a \cdot f = f \text{ for all } a \in G\},$$

and let

$$J^+ = \{f \in J \mid f(0) = 0\}.$$

We will often be concerned with the homogeneous ideal J^+S in S generated by J^+.

PROPOSITION 1. *Let L be any graded subspace of S such that $S = J^+S + L$ is a direct sum. Then*

$$S = JL.$$

Proof. We must show $S^k \subseteq JL$ for all k. This is obvious if $k = 0$ since $S^0 \subseteq L$. Assume it is true for S^j where $j \leqq k$. But since one clearly has $S^{k+1} \subseteq J^+S^{(k)} + L$ where $S^{(k)} = \sum_{i=0}^{k} S^i$ it is then obviously true for S^{k+1}. Q. E. D.

We are interested first of all in the question as to when S is free over J, or more specifically, as to when S is a tensor product of J and L. (Choosing L to be G-stable, such a decomposition of S reduces the study of its G-module structure to that of L.) We first observe that the two conditions are equivalent.

The expression linear independence (resp. basis) without any reference

to a ring always means linear independence (resp. basis) with respect to $\boldsymbol{C}$. Furthermore, tensor product without reference to a ring means tensor product over $\boldsymbol{C}$.

LEMMA 1. *The following conditions are equivalent:*

1. *Let L be as in Proposition 1. Then the map*

$$J \otimes L \to S$$

given by $f \otimes g \to fg$ is an isomorphism.

2. *S is free over J.*

3. *Let $M \subseteq S$ be any subspace such that $M \cap J^+S = (0)$. Then for elements in M linear independence is equivalent to linear independence over J.*

Proof. Obviously $(1) \Rightarrow (2)$ since a $\boldsymbol{C}$ basis of L defines a J basis of S. Assume (2) and let e_i, $i = 1, 2, \cdots$, define a J basis of S. Let $f_1, \cdots, f_k$ be linearly independent elements of M. For $j = 1, \cdots, k$ write $f_j = \sum_i a_{ij} e_i$ where we may assume $i = 1, 2, \cdots, p$ and $a_{ij} \in J$. To show that the f_j are linearly independent over J it clearly suffices to see that the $k \times p$ matrix (a_{ij}) with entries in J is of rank k. If f' denotes the image of $f \in S$ in S/J^+S it is clear that $f'_j = \sum a_{ij}(0) e'_i$. Since the f'_j are obviously linearly independent it is clear that the $\boldsymbol{C}$ matrix $a_{ij}(0)$ is of rank k. Hence not all $k \times k$ minors of the J matrix (a_{ij}) can be zero. This proves that the f_j are linearly independent over J.

It is trivial that linear independence over J implies linear independence.

To obtain (1) from (3) we simply put $M = L$ and apply Proposition 1.

Q. E. D.

1.2. Now for any $x \in X$ let

$$O_x = \{y \in X \mid y = ax, \text{ for some } a \in G\}.$$

A subset $O \subseteq X$ is said to be an orbit if $O = O_x$ for some $x \in X$.

For any $x \in X$ let

$$G^x = \{a \in G \mid ax = x\}$$

It is clear that G^x is an algebraic (hence a closed, complex Lie subgroup) subgroup of G. Furthermore if

$$\beta'_x \colon G \to X \tag{1.2.1}$$

is the map given by $\beta'_x(a) = ax$ then β'_x induces a bijection

$$\beta_x \colon G/G^x \to O_x \tag{1.2.2}$$

Now let U be universal enveloping algebra of the Lie algebra of G. Then since the representation of G on S induces a representation of its Lie algebra on S it is clear that S becomes a U-module by further extension to U. Obviously S^k is a U-submodule for any k. We denote by $p \cdot f \in S$ the effect of applying p to f where $p \in U$ and $f \in S$.

For any subset $P \subseteq X$ and $f \in S$ let $f \mid P$ be the restriction of f to P.

LEMMA 2. *Let $f_j \in S$, $j = 1, 2, \cdots, k$, and let p_i, $i = 1, 2, \cdots$, be a basis of U. Consider the k-column matrix $D = (d_{ij})$ where $d_{ij} = p_i \cdot f_j$. (The matrix thus has entries in S and hence, for any $x \in X$, $D(x) = (d_{ij}(x))$ is a $\boldsymbol{C}$ matrix.)*

Then if $x \in X$ the functions $f_i \mid O_x$ on O_x, $j = 1, 2, \cdots, k$, are linearly independent if and only if $D(x)$ has rank k.

Proof. Let C be the algebra of all holomorphic functions on the complex homogeneous space G/G^x. It is clear that C is a module for the Lie algebra of G since the latter defines holomorphic vector fields on G/G^x. Hence C also becomes a U-module and in such a way that if

$$\alpha : S \to C$$

is the homomorphism given by $\alpha f = f \circ \beta_x$ one has

$$\alpha(p \cdot f) = p \cdot \alpha f \tag{1.2.3}$$

for any $f \in S$. Furthermore if $s \in G/G^x$ denotes the point corresponding to G^x and $g \in C$ then since the Lie algebra of G spans the holomorphic tangent space at s it follows from the Taylor expansion that g vanishes identically on G/G^x if and only if $(p \cdot g)(s) = 0$ for all $p \in U$.

Now assume that rank $D(x) < k$. Then there exists a non-zero vector $(c_1, \cdots, c_k) \in \boldsymbol{C}^k$ such that $\sum_j d_{ij}(x) c_j = 0$ for all i. Thus if $f = \sum_j c_j f_j \in S$ one has $(p \cdot f)(x) = 0$ for all $p \in U$. Thus by (1.2.3) $(p \cdot \alpha f)(s) = 0$ for all $p \in U$ and hence αf vanishes identically on G/G^x; or equivalently $f \mid O_x$ is zero. Hence the $f_j \mid O_x$ are linearly dependent. Conversely assume that the $f_j \mid O_x$ are linearly dependent so that $(\sum c_j f_j) \mid O_x$ is zero for a non-zero vector $(c_1, \cdots, c_k) \in \boldsymbol{C}^k$. But then $\alpha(\sum_j c_j f_j) = 0$ and hence for $i = 1, 2, \cdots$,

$$\begin{aligned} \sum_j c_j p_i \cdot f_j(x) &= (p_i \cdot \alpha(\sum_j c_j f_j))(s) \\ &= 0 \end{aligned}$$

Thus rank $D(x) < k$. Q. E. D.

As a corollary we obtain the following criterion for linear independence over J.

LEMMA 3. *Let $f_j \in S$, $i = 1, 2, \cdots, k$. Assume the functions $f_j \mid O$ are linearly independent for some orbit $O \subseteq X$. Then the f_j are linearly independent over J.*

Proof. Assume $\sum_j g_j f_j = 0$ where $g_j \in J$, $j = 1, 2, \cdots, k$. Let D be the matrix given in Lemma 2 and let $x \in O$. Then by Lemma 2 there exists a $k \times k$ minor of D whose determinant $\alpha \in S$ is such that $\alpha(x) \neq 0$. But then there exists a neighborhood W of x such that $\alpha(y) \neq 0$ for all $y \in W$. Thus $D(y)$ has rank k and hence the $f_j \mid O_y$ are linearly independent for all $y \in W$. But since the g_j reduce to constants on any orbit O_y it follows from the relation $\sum_j g_j f_j = 0$ that the g_j vanish identically on W. This implies that the g_j vanish on X since the g_j are polynomials. Q. E. D.

1.3. For any subset $Y \subseteq X$ we will let

$$I(Y) = \{f \in S \mid f \mid Y = 0\}$$

be the ideal in S defined by Y.

Now let $P \subseteq X$ be the cone (since J^+ is homogeneous) given by

$$P = \{x \in X \mid f(x) = 0 \text{ for all } f \in J^+\}$$

Since P is defined by the ideal J^+S one knows that $I(P)$ is the radical of J^+S and that the cone P is irreducible (in the sense of algebraic geometry) and $J^+S = I(P)$ if and only if J^+S is prime.

We now give a criterion for the conditions of Lemma 1 to be satisfied.

PROPOSITION 2. *Assume* (1) *J^+S is prime and* (2) *there exists an orbit O such that $\bar{O} = P$. Then the conditions of Lemma* 1, § 1.1, *are satisfied. In particular if $S = J^+S + L$ is a direct sum where L is a graded subspace of S then the map*

$$J \otimes L \to S \tag{1.3.1}$$

given by $f \otimes g \to fg$, is an isomorphism.

Proof. Let $M \subseteq S$ be any subspace such that $M \cap J^+S = (0)$. Since $J^+S = I(P)$ it is clear that if $f_j \in M$, $j = 1, 2, \cdots, k$, are linearly independent then the $f_j \mid P$ are linearly independent. But this obviously implies that the $f_j \mid O$ are linearly independent since $\bar{O} = P$. But then the f_j are linearly independent over J by Lemma 3 and thus the result follows by (1) and (3) of Lemma 1. Q. E. D.

Remark 1. In the proof we have only used the fact that $J^+S = I(P)$ and not that J^+S is prime. However assumption (2) already implies that the cone P is irreducible (recall that G is connected) so that there is no loss in assuming that J^+S is prime.

1.4. Now assume that B is a symmetric non-singular G-invariant bilinear form on X. Then, as one knows, B induces a unique G-module ring isomorphism (also written B)

$$B : S_* \to S \tag{1.4.1}$$

of degree zero where $\langle \partial_x, B\partial_y \rangle = B(x, y)$ for any $x, y \in X$. Obviously for any $\partial_1, \partial_2 \in S_*$ one has

$$\langle \partial_1, B\partial_2 \rangle = \langle \partial_2, B\partial_1 \rangle. \tag{1.4.2}$$

Now let

$$J_* = \{\partial \in S_* \mid a \cdot \partial = \partial \text{ for all } a \in G\}$$

and let J_*^+ be the space of elements in J_* having zero constant term. It is obvious that

$$B(J_*) = J \text{ and } B(J_*^+) = J^+.$$

An element $f \in S$ is called G-harmonic in case

$$\partial f = 0$$

for all $\partial \in J_*^+$. Let H be the (obviously graded) space of all G-harmonic polynomials in S. By (1.1.2) and (1.1.4) it is clear that H is a G-submodule of S.

One obtains a class of G-harmonic polynomials in the following way: Let P' be the cone in S^1 defined by putting

$$P' = B\{\partial_x \in S_1 \mid x \in P\}$$

and let $H_P \subseteq S$ be the (graded) space spanned by all powers z^m, $m = 0, 1, \cdots$, for all $z \in P'$. The following proposition was proved independently by Helgason (see [10]).

PROPOSITION 3. *One has $H_P \subseteq H$. Furthermore $S = J^+S + H$ is a G-module direct sum so that by Proposition* 1, § 1.1,

$$S = JH.$$

Proof. Let $z \in P'$ so that $z = B(\partial_x)$ for $x \in P$. Let $\partial \in J_k$ where $k > 0$. We wish to show that $\partial z^m = 0$. We may assume that $m \geqq k$ and hence by

(1.1.2) it suffices to show that $\langle \partial\partial_1, z^m \rangle = 0$ for all $\partial_1 \in S_{m-k}$. But by (1.4.2)

$$\begin{aligned}\langle \partial\partial_1, z^m \rangle &= \langle \partial\partial_1, B(\partial_x{}^m) \rangle \\ &= \langle \partial_x{}^m, B(\partial\partial_1) \rangle \\ &= \langle \partial_x{}^m, ff_1 \rangle\end{aligned}$$

where $f = B\partial$ and $f_1 = B\partial_1$. But $f \in J^k$ and hence $ff_1 = g \in J^+S$. Thus $\langle \partial\partial_1, z^m \rangle = \langle \partial_x{}^m, g \rangle = m!g(x) = 0$ since $x \in P$ and $g \in I(P)$. Hence $H_P \subseteq H$.

Now let $H' \subseteq S$ be the orthocomplement of $J_*{}^+S_*$ in S. Thus if $f \in S$ then $f \in H'$ if and only if

$$\begin{aligned}\langle \partial\partial_1, f \rangle &= \langle \partial_1, \partial f \rangle \\ &= 0\end{aligned}$$

for all $\partial_1 \in S$ and $\partial \in J_*{}^+$. It follows immediately then that $H' = H$. To prove that $S = J^+S + H$ is a direct sum therefore it suffices, by dimension, (one restricts to S^k) to show that $H \cap J^+S = 0$. (One already uses that $B(J_*{}^+S_*) = J^+S$). But to show $H \cap J^+S = 0$ it suffices to show that B induces a non-singular bilinear form on $J_*{}^+S_*$. Now let K be a maximal compact subgroup of G. Then one knows there exists a real subspace $X_R \subseteq X$ such that (1) X_R is stable under K, (2) B is positive definite on X_R and (3) $X = X_R + iX_R$ is a real direct sum. But by (2) it is obvious that B induces a positive definite bilinear form on $S_{*,R}(X_R) = S_R$. But by (1) and (3) $J_*{}^+$ is the complexification of its intersection with S_R since G is the complexification of K. Hence $J_*{}^+S_*$ is also the complexification of its intersection with S_R and consequently B induces a non-singular bilinear form on $J_*{}^+S_*$. Q. E. D.

Now combining Propositions 2, § 1.3, and 3, § 1.4, we obtain the following "separation of variables" theorem.

PROPOSITION 4. *Assume that G leaves invariant a symmetric non-singular bilinear form on X.*

Let S, H and J be, respectively, the ring of all polynomials on X, the space of G-harmonic polynomials on X and the ring of invariant polynomials on X.

Let $P \subseteq X$ be the homogeneous affine variety defined by the ideal J^+S in S.

Now assume (1) *that there exists an orbit O such that $\bar{O} = P$ and* (2) *the ideal J^+S is prime. Then the mapping*

$$J \otimes H \rightarrow S \tag{1.4.3}$$

given by $f \otimes g \to fg$ is a G-module isomorphism. Furthermore if $P' \subseteq S^1$ is the cone corresponding (by means of the bilinear form) to P and H_P is the subspace of S generated by all powers z^m, $z \in P'$, $m = 0, 1, \cdots$, then

$$H = H_P. \tag{1.4.4}$$

Proof. Since G operates as algebra automorphisms of S it is obvious that (1.4.3) is a G-module map. But it is also an isomorphism by Propositions 2 and 3. Now let $H_* \subseteq S_*$ be the space generated by all powers ∂_x^m where $x \in P$. Let B be as in (1.4.1). Then clearly $B(H_*) = H_P$. Now if $H_P^m \neq H^m$ then by dimension and Proposition 3 there exist a non-zero $f \in H^m$ such that $\langle \partial, f \rangle = 0$ for all $\partial \in H_m$. But then putting $\partial = \partial_x^m/m!$ for $x \in P$ it follows from (1.1.3) that $f(x) = 0$ for all $x \in P$. But then $f \in J^+S$ since J^+S is prime. This is a contradiction since $J^+S \cap H = (0)$. Q. E. D.

Remark 2. A familiar instance as to when the conclusion of Proposition 4 holds is the case where $X = \boldsymbol{C}^n$, $n \geqq 2$, and G is the full complex rotation group. Here J is the algebra generated by 1 and $z_1^2 + \cdots + z_n^2$, H is space of all polynomials which satisfy Laplace's equation.

$$\sum_{i=1}^{n} \frac{\partial^2 f}{\partial z_i^2} = 0$$

and P is the conic given by $z_1^2 + \cdots + z_n^2 = 0$. If $n \geqq 3$ one obtains the classical separation of variables theorem as a consequence of Proposition 4 since (1) $P = \bar{O}$, where O is the set (easily seen to be an orbit) of all vectors $x \in \boldsymbol{C}^n$ such that $x \in P$ and $x \neq 0$, and (2) J^+S is a prime ideal, because it is generated by $z_1^2 + \cdots + z_n^2$ and this polynomial is irreducible if $n \geqq 3$.

1.5. For any element $x \in X$ besides O_x, we may consider the orbit O_{cx} where $c \in \boldsymbol{C}^*$ is any non-zero scalar. It is obvious of course that $O_{cx} = cO_x$. Now where $P \subseteq X$ is the cone defined in § 1.3 and $x \in X$ is arbitrary put

$$P_x = \overline{\bigcup_{c \in \boldsymbol{C}^*} O_{cx}} \cap P$$

It is clear of course that $P_x \subseteq P$ is stable under the action of G. An element $x \in X$ will be called quasi-regular if

$$P_x = P$$

Remark 3. Note that if $P = \bar{O}$ for an orbit O then any element of O is quasi-regular.

Now for any subset $W \subseteq X$ let $S(W)$ be the ring of all functions on W of the form $g \mid W$ where $g \in S$. Note that if W is stable under G then $S(W)$ is a G-module with respect to the action of G given by (1.1.5) where $f \in S(W)$.

Let $x \in X$. We are particularly interested in the ring $S(O_x)$ of functions on the orbit O_x.

Obviously the map

$$S \to S(O_x) \tag{1.5.1}$$

defined by the correspondence $f \to f \mid O_x$ is a G-module epimorphism.

PROPOSITION 5. *As in Proposition* 1, § 1.1, *let* $L \subseteq S$ *be any graded subspace such that* $S = J^+S + L$ *is a direct sum. Also for any* $x \in X$ *let*

$$\gamma_x \colon L \to S(O_x)$$

be the linear map obtained by restricting (1.5.1) *to* L. *Then* γ_x *is an epimorphism.*

Assume that conditions (1) *and* (2) *of Proposition* 2, § 1.3, *are satisfied. Then* γ_x *is an isomorphism for any quasi-regular element* $x \in X$.

In particular if G *leaves invariant a non-singular symmetric bilinear form on* X *and, as in Proposition* 4, § 1.4, $L = H$ *is the space of* G*-harmonic polynomials on* X *then*

$$\gamma_x \colon H \to S(O_x) \tag{1.5.2}$$

is a G*-module isomorphism for any quasi-regular element* $x \in X$.

Proof. Since J reduces to scalars on any orbit O_x it follows from Proposition 1 that γ_x maps L surjectively onto $S(O_x)$. Assume now that conditions (1) and (2) of Proposition 2 are satisfied. Let x be quasi-regular. We must show that γ_x is injective. Let $f \in L$. Since L is graded we may write $f = \sum_{i=1}^{k} c_i f_i$ where $f_i \in L$ is homogeneous of degree n_i, and the f_i, $i = 1, 2, \cdots, k$, are linearly independent.

But now since J^+S is the prime ideal corresponding to P it follows that the functions $f_i \mid P$ are linearly independent. But since $P = \bar{O}$ for an orbit O one has also that the functions $f_i \mid O$ are linearly independent. The argument of Lemma 2, § 1.2, shows that for any $y \in O$ there exists a neighborhood W of y in X such that for any $z \in W$ the functions $f_i \mid O_z$ are linearly independent. But now since $y \in P = P_x$ there exists a non-zero scalar c such that $O_{cx} = O_z$ for some $z \in W$. Hence there exists a non-zero scalar c such that the functions $f_i \mid O_{cx}$ are linearly independent. Now let

$$\mu \colon O_{cx} \to O_x$$

be the bijection defined by $y \to 1/c \cdot y$. If μ^* is then the corresponding contravariant isomorphism on functions one has

$$\mu^*(f_i \mid O_x) = 1/c^{n_i}(f_i \mid O_{cx})$$

But then since the $f_i \mid O_{cx}$ or $1/c^{n_i}(f_i \mid O_{cx})$ are linearly independent it follows that the $f_i \mid O_x$ are linearly independent. But then if $f \mid O_x$ is zero it follows that the c_i are all zero and hence f is identically zero. Thus γ_x is injective.

The isomorphism (1.5.2) is a G-module map since (1.5.1) is a G-module map. Q.E.D.

Remark 4. In the example of Remark 2, § 1.4, not that $x \in \boldsymbol{C}^n$ is quasi-regular if and only if $x \neq 0$.

Thus in that example one has that

$$\gamma_x \colon H \to S(O_x) \tag{1.5.3}$$

is an isomorphism for any $x \neq 0$. Hence all $S(O_x)$ where $x \neq 0$ are equivalent as G-modules.

1.6. In order to apply Propositions 3, 4 in § 1.4 or Proposition 5, § 1.5, one needs to know that J^+S is a prime ideal. In general this appears to be difficult to ascertain even if one knows J completely. (Except of course if J has only one ring generator, as in the example of Remark 2.) However we will now observe (Proposition 6, § 1.6) that in the familiar case when J is a polynomial ring the question of the primeness of J^+S reduces to a more manageable one.

Throughout much of the remainder of the paper we will need to draw upon techniques and results in algebraic geometry. Our reference for all definitions will be [3] where for us the fixed algebraically closed field is of course $\boldsymbol{C}$. We recall in particular that by definition, among other things, a variety is irreducible in its Zariski topology.

To avoid confusion of terminology we remark here that the words open, closed, closure and denseness, etc. will have their usual Hausdorff topological meaning unless stated otherwise (i.e., unless preceded by "Zariski").

If $f_i \in S$, $i = 1, 2, \cdots, l$, are arbitrary let $(f_1, \cdots, f_l)$ denote the ideal in S that they generate. If $Y \subseteq X$ is a Zariski closed subvariety of X of dimension $n - l$ then we recall that Y is called a complete intersection in case

$$I(Y) = (f_1, \cdots, f_l)$$

for some $f_i \in I(Y)$, $i = 1, \cdots, l$.

Now for any $f \in S$ and $x \in X$ let $(df)_x$ be the value of the differential df

at x. If $f_i \in S$, $i = 1, \cdots, l$, then one knows that the $(df_i)_x$ are linearly independent if and only if the $n \times l$ matrix $(\partial_{x_j} f_i)(x)$, $j = 1, \cdots, n$, has rank l where the x_j is any basis of X.

The following lemma in one form or another is well known in algebraic geometry.

LEMMA 4. *Let $f_i \in S$, $i = 1, 2, \cdots, l$, and let Y be the Zariski closed set given by*

$$Y = \{x \in X \mid f_i(x) = 0, i = 1, \cdots, l\}.$$

Assume (1) *Y is a subvariety of X (that is, assume Y is irreducible) and* (2) *there exists $y \in Y$ such that $(df_i)_y$, $i = 1, 2, \cdots, l$, are linearly independent. Then Y is a subvariety of* $\dim n - l$. *Furthermore*

$$I(Y) = (f_1, \cdots, f_l) \tag{1.6.1}$$

so that (a) *$(f_1, \cdots, f_l)$ is a prime ideal and* (b) *Y is a complete intersection.*

Proof. Let S_y be the local ring of X at y. Let $I = (f_1, \cdots, f_l)$. Since the $(df_i)_y$ are linearly independent the f_i may be included in a complete system of uniformizing variables at y. Thus by [3], Proposition 3, p. 219, IS_y is a prime ideal of S_y. Furthermore since $I(Y)$ is the radical of I in S it is clear that IS_y is the ideal of Y at y (that is, $I(Y)S_y = IS_y$) so that, by the same reference, $\dim Y = n - l$ and $IS_y \cap S = I(Y)$. To prove (1.6.1) it suffices to show that I is primary for $I(Y)$ since in that case $IS_y \cap S = I$ (a primary ideal is equal to the contraction of its extension; see [13], Theorem 19, p. 228). But I is primary by MacCaulay's theorem (see [19], p. 203) which asserts that there are no embedded primes for I so that $I(Y)$ is the only associated prime ideal. Q.E.D.

Now we recall that G is a connected algebraic reductive group. Hence G has the structure of an affine variety. (It is Zariski closed in $\operatorname{Aut} X$ but not necessarily Zariski closed in $\operatorname{End} X$.) Since (1.2.1) is obviously a morphism it follows that any orbit $O \subseteq X$ is an irreducible constructible set. In fact since O is epais ([3], Proposition 4, p. 95) and G operates transitively on it, it follows that O is a subvariety ([3], Theorem 5, p. 68) of X. It follows therefore that its (usual) closure $\bar{O}$ is a Zariski closed subvariety of the same dimension as O.

As an application of Lemma 4, § 1.6, we have

PROPOSITION 6. *Assume J, as a ring, is generated by l homogeneous algebraically independent polynomials u_i, $i = 1, 2, \cdots, l$.*

Now let $\xi \in \mathbf{C}^l$, $\xi = (\xi_1, \cdots, \xi_l)$, *be an arbitrary complex l-tuple and let*

$$P(\xi) = \{x \in X \mid u_i(x) = \xi_i, i = 1, 2, \cdots, l\}$$

Assume $P(\xi)$ *is not empty and there exists an orbit* $O(\xi)$ *such that*

$$P(\xi) = \overline{O(\xi)}. \tag{1.6.2}$$

Then $P(\xi)$ *is a Zariski closed subvariety (of* X*) of dimension* $n - l$. *Furthermore the ideal* $(u_1 - \xi_1, \cdots, u_l - \xi_l)$ *in* S *is prime if and only if there exists* $y \in P(\xi)$ *such that the* $(du_i)_y$ *are linearly independent. In such a case* $P(\xi)$ *is a complete intersection and the set* $P(\xi)_s$ *of simple points on* $P(\xi)$ *is given by*

$$P(\xi)_s = \{x \in P(\xi) \mid (du_i)_x, i = 1, 2, \cdots, l, \text{ are linearly independent}\}. \tag{1.6.3}$$

Proof. Since $O(\xi)$ is irreducible it follows from (1.6.2) that $P(\xi)$ is a subvariety of the same dimension as $O(\xi)$. Now by [3], Corollary, p. 102, it is clear that

$$\dim P(\xi) \geqq n - l.$$

To prove that $\dim P(\xi) = n - l$ it suffices to show that

$$\dim O \leqq n - l \tag{1.6.4}$$

for any orbit O.

Let m be the maximum of the dimensions of all orbits. Let $\mathfrak{u} \subseteq \operatorname{End} X$ be the Lie algebra of G and for each $x \in X$ let ϕ_x be the homomorphism of $\mathfrak{u}$ into X given by $\phi_x(z) = z(x)$. It is then obvious that $\operatorname{rank} \phi_x = \dim O_x$. By consideration of minors it is then clear that

$$U = \{x \in X \mid \dim O_x = m\}$$

is a non-empty Zariski open subset of X. Now let V be the Zariski open subset of X consisting of all $x \in X$ such that the $(du_i)_x$ are linearly independent. To see that $V \cap U$ is not empty it clearly suffices to see that V is not empty. But this is a known consequence of algebraic independence. Indeed if $u_j \in S$, $j = l + 1, \cdots, n$, are chosen so that u_i, $i = 1, 2, \cdots, n$, is a transcendental basis of S then each element of a coordinate basis z_j, $j = 1, 2, \cdots, n$ of X is algebraically dependent upon the u_i. Hence on a non-empty Zariski open set each dz_j is in the span of the du_i. This proves that V and hence $V \cap U$ is not empty. Now let $x \in V \cap U$ and let W be the $n - l$ dimensional variety (see [3], Proposition 3, p. 219) containing x whose prime ideal at x is IS_x where $I = (u_1 - u_1(x), \cdots, u_l - u_l(x))$ and S_x is the local ring at x.

Since, obviously, $O_x \subseteq W$ and $\dim O_x = m$ it follows that $m \leqq n - l$ and this proves (1.6.4) and hence $\dim P(\xi) = n - l$.

Now if the ideal $(u_1 - \xi_1, \cdots, u_l - \xi_l)$ is prime it must equal $I(P(\xi))$ and hence $(df)_x$ for any $x \in P(\xi)$ and $f \in I(P(\xi))$ lies in the span of the $(d(u_i - \xi_i))_x = (du_i)_x$. But since $\dim P(\xi) = n - l$ one immediately obtains (1.6.3) by Zariski's criterion and since $(P(\xi))_s$ is not empty there exists $y \in P(\xi)$ such that the $(du_i)_y$ are linearly independent. Conversely if the latter holds $(u_1 - \xi_1, \cdots, u_l - \xi_l)$ is prime by Lemma 4, § 1.6, and hence $P(\xi)$ is a complete intersection. Q. E. D.

For us Proposition 4, § 1.4, will be put into effect by

PROPOSITION 7. *Assume J, as a ring, is generated by l algebraically independent homogeneous polynomials u_i, $i = 1, 2, \cdots, l$.*

Assume also that there exists an orbit O such that $P = \bar{O}$. Then P is a subvariety of dimension $n - l$. Moreover J^+S is prime if and only if there exists $y \in P$ such that $(du_i)_y$, $i = 1, 2, \cdots, l$, are linearly independent.

Proof. This is just the special case $\xi = 0$ of Proposition 6. Q. E. D.

2. Normality and the closure of an orbit. 1. If Y is any variety we let $R(Y)$ denote the ring of everywhere defined rational functions on Y.

Now let C be the ring of all holomorphic functions on G. If $f \in C$, $a \in G$, then the left (resp. right) translate $a \cdot f$ (resp. $f \cdot a$) of f by a is the function defined by putting $(a \cdot f)(b) = f(a^{-1}b)$ (resp. $(f \cdot a)(b) = f(ba^{-1})$). It is obvious that $a \cdot f \in C$ (resp. $f \cdot a \in C$) for all $f \in C$, $a \in G$.

One knows that $R(G)$ is a subring of C which in fact may be given by (see [9], Theorem 5.2)

$R(G) = \{f \in C \mid \text{space spanned by all } a \cdot f, a \in G, \text{ is finite dimensional}\}$
It is obvious that $R(G)$ is stable under left and right translations.

Now let D denote the set of equivalence classes of all irreducible rational (equivalently, holomorphic) finite dimensional representations of G. For each $\lambda \in D$ choose a fixed irreducible representation.

$$\nu^\lambda \colon G \to \operatorname{Aut} V^\lambda$$

belonging to λ. The dual space to V^λ will be denoted by V_λ and the irreducible representation of G on V_λ contragradient to ν^λ will be denoted by ν_λ.

If M is any G-module we will let M^λ denote the set of all vectors in M which transform according to the irreducible representation ν^λ. Since G is assumed to be reductive one knows that if each vector in M generates a finite

dimensional cyclic G-module then M is in fact a direct sum of the M^λ. In particular regarding $R(G)$ as a G-module under left translation one has that

$$R(G) = \sum_{\lambda\in D} R^\lambda(G)$$

is a direct sum. Since $\nu^\lambda(G)$ generates End V^λ one can be very explicit about the structure of $R^\lambda(G)$. In fact let d_λ be the dimension of V_λ and let v_i and v'_j, $i, j = 1, 2, \cdots, d_\lambda$, be, respectively, a basis of V^λ and a basis of its dual space V_λ. Now let $g_{ij}{}^\lambda$ be the function on G defined by

$$(2.1.1) \qquad g_{ij}{}^\lambda(b) = \langle v_i, \nu_\lambda(b)v'_j\rangle.$$

Then one knows that the $d_\lambda{}^2$ functions defined in this way form a basis of $R^\lambda(G)$. In particular $R^\lambda(G)$ is finite dimensional and in fact

$$\dim R^\lambda(G) = d_\lambda{}^2.$$

Now assume that $F \subseteq G$ is an algebraic (and hence closed, Lie) subgroup. Then one knows that G/F (space of left coset $aF, a \in G$) has the structure of an irreducible algebraic variety where

$$\dim G/F = \dim G - \dim F$$

and the ring $R(G/F)$ of everywhere defined rational functions on G/F may be identified, in the obvious way, with the set of elements in $R(G)$ that are right invariant under F.

Now for any $\lambda \in D$ let $V_\lambda{}^F$ be the space of all vectors in the dual space V_λ to V^λ that are fixed under all transformations on V_λ of the form $\nu_\lambda(a)$ where $a \in F$. Put

$$d_\lambda{}^F = \dim V_\lambda{}^F.$$

Now it is obvious that $R(G/F)$ is a G-submodule (by left translations) of $R(G)$. It is furthermore obvious that

$$R^\lambda(G/F) \subseteq R^\lambda(G).$$

The following is a special case of an algebraic Frobenius reciprocity theorem. We prove for it for completeness.

PROPOSITION 8. *For $i = 1, \cdots, d_\lambda$ and $j = 1, \cdots, d_\lambda{}^F$ let v_i be a basis of V^λ and let w'_j be a basis of $V_\lambda{}^F$. Also let h_{ij} be the function on G given by*

$$h_{ij}(b) = \langle v_i, \nu_\lambda(b)w'_j\rangle.$$

Then $h_{ij} \in R^\lambda(G/F)$ and in fact the $d_\lambda d_\lambda{}^F$ functions defined in this way are a basis of $R^\lambda(G/F)$.

Thus

$$\dim R^\lambda(G/F) = d_\lambda d_\lambda^F$$

so that ν^λ *occurs with multiplicity* d_λ^F *in* $R(G/F)$. *Furthermore*

$$R(G/F) = \sum_{\lambda \in D} R^\lambda(G/F) \tag{2.1.2}$$

is a direct sum.

Proof. The decomposition (2.1.2) is obvious since each element of $R(G)$ is an element of $R(G)$ and hence generates a finite dimensional subspace under the action (left translation) of G.

Furthermore it is also obvious that the $d_\lambda d_\lambda^F$ functions h_{ij} defined in the proposition are in $R^\lambda(G/F)$ and (see (2.1.1)) are linearly independent. To prove the proposition therefore one simply has to show that every element of $R^\lambda(G)$ invariant under right translation by elements of F is in the span of the h_{ij}. Assume that $g \in R^\lambda(G)$ and $g \cdot a = g$ for all $a \in F$. Let g_{ij}^λ be as in (2.1.1) (a basis of $R^\lambda(G)$). Write $g = \sum g_{ij}^\lambda c_{ji}$ where $c_{ij} \in \boldsymbol{C}$ defines a matrix and hence, relative to the basis v'_j, a linear transformation α of V_λ. It suffices only to show that $\operatorname{Im} \alpha \subseteq V_\lambda^F$. But the condition on g implies that $(\nu_\lambda(a) - 1)\alpha = 0$ for all $a \in F$. This proves $\operatorname{Im} \alpha \subseteq V_\lambda^F$. Q.E.D.

Remark 5. A case of importance for us is the case where $F = A$ is a Cartan subgroup of G. Here V_λ^A is just the zero weight subspace, corresponding to A, of V_λ. To make it independent of A we will put $l_\lambda = d_\lambda^A$ so that $l_\lambda =$ multiplicity of the zero weight of ν_λ (2.1.3).

Remark 6. Since one knows that the multiplicity of any weight μ for ν_λ is equal to the multiplicity of $-\mu$ for ν^λ it follows that l_λ is also the multiplicity of the zero weight of ν^λ.

2.2. Now we wish to apply the considerations of § 2.1 to the case where $F = G^x$ for any $x \in X$. See § 1.2. By Proposition 8 any question as to the complete reduction of $R(G/G^x)$ as a G-module becomes a question in the finite dimensional representation theory of G and how such representations restrict to G^x.

Now, as we observed in § 1.6, the orbit O_x is a subvariety of X. Furthermore the bijection $\beta_x\colon G/G^x \to O_x$ induced by β'_x is an algebraic isomorphism (this follows easily from the transitivity of G together with [3], Corollary, p. 53 and Corollary 2, p. 90. (See also [1], § 2.2.) Thus if $R(O_x)$ is regarded as a G-module, using the action of G in O_x, it follows that β_x induces a G-module and ring isomorphism

$$R(G/G^x) \to R(O_x). \tag{2.2.1}$$

Now we recall that $S(O_x)$ is the ring of functions on O_x obtained by restricting S (the ring of polynomials on X) to O_x. Since β_x is a morphism one obviously has

$$S(O_x) \subseteq R(O_x)$$

for any $x \in X$ and in fact it is clear that $S(O_x)$ is a G-submodule of $R(O_x)$. Unlike $R(O_x)$ whose G-module structure is completely determined by Proposition 8 because (2.2.1) is a G-module isomorphism, in the general case it seems (to us) to be quite difficult to describe how $S(O_x)$ decomposes as a G-module.

In many instances, however, $S(O_x) = R(O_x)$ (and hence, in such cases, one knows the G-module structure of $S(O_x)$). Indeed, in the general case since $\bar{O}_x$ is Zariski closed in X one has

$$S(\bar{O}_x) = R(\bar{O}_x). \tag{2.2.2}$$

Thus

$$S(O_x) = R(O_x) \text{ if } O_x = \bar{O}_x. \tag{2.2.3}$$

Remark 7. In the example of Remark 2, § 1.4, one depends upon the equality $S(O_x) = R(O_x)$ for a particular x in order to solve the Dirichlet problem in $\boldsymbol{R}^n$. Indeed let $x \in \boldsymbol{R}^n$ where $(x,x) = \alpha > 0$ and let f be a continuous function on the sphere $S^{n-1} = O_x \cap \boldsymbol{R}^n$ of radius $\sqrt{\alpha}$. The problem is to extend f as a harmonic function f' defined in the interior of S^{n-1}. To do this one expands f

$$f = \sum_{\lambda \in D} c_\lambda f_\lambda,$$

using some limiting process (e.g., L_2), as an infinite sum of spherical harmonics f_λ. That is, here $c_\lambda \in \boldsymbol{C}$ and $f_\lambda = g_\lambda \mid S^{n-1}$ where

$$g_\lambda \in R^\lambda(O_x)$$

However since $R(O_x) = S(O_x)$ it follows from (1.5.3) that there exists a unique harmonic polynomial $h_\lambda \in H$ on $\boldsymbol{C}^n$ such that $h_\lambda \mid O_x = g_\lambda$. One then puts

$$f' = \sum_{\lambda \in D} c_\lambda h'_\lambda$$

where h'_λ is the restriction of h_λ to the interior of S^{n-1}.

Now it is not necessarily true, in general, that $S(O_x) = R(O_x)$. For example let X be the m^2 dimensional space of all complex $m \times m$ matrices and G is the general linear group $Gl(m, \boldsymbol{C})$ regarded as operating on X by left matrix multiplication. Then if x is the identity matrix O_x is isomorphic to

$Gl(m, \boldsymbol{C})$. But $S(O_x) \neq R(O_x)$ since in particular if $f(a) = (\det a)^{-1}$ for $a \in G$ then $f \in R(O_x)$ but $f \notin S(O_x)$.

The equality $S(O_x) = R(O_x)$ in the example of Remark 7 when $(x, x) > 0$ may be established either using the fact that O_x is closed (see (2.2.3)) or by applying the Stone-Weierstrass theorem to both $S(O_x)$ and $R(O_x)$ restricted to $O_x \cap \boldsymbol{R}^n$. These methods also work more generally in case $(x, x) \neq 0$. However, they do not apply to O_x where $x \neq 0$ and $(x, x) = 0$. Nevertheless it is still true in this case that $R(O_x) = S(O_x)$. The more powerful tool (and the one that will be required in § 5.1) needed to establish the equality for this case is given in the next proposition.

For any $x \in X$ let C_x be the Zariski closed subset of X defined by taking the complement of O_x in $\bar{O}_x$. If we put

$$\operatorname{codim} C_x = \dim \bar{O}_x - \dim C_x$$

then of course one has

$$\operatorname{codim} C_x \geqq 1.$$

An affine variety Y is called normal in case the ring $R(Y)$ is integrally closed in its quotient field.

PROPOSITION 9. *Let* $x \in X$. *Assume* (1) *that* $\bar{O}_x$ *is a normal variety and* (2) *that* $\operatorname{codim} C_x \geqq 2$. *Then*

$$S(O_x) = R(O_x).$$

Proof. If Y is any variety let $Q(Y)$ denote the field of all rational functions on Y. In any $f \in Q(O_x)$ let $\bar{f}$ denote its image in $Q(\bar{O}_x)$ under the canonical isomorphism $Q(O_x) \to Q(\bar{O}_x)$ defined by extension.

Now let $f \in R(O_x)$. Then obviously $\bar{f} \in Q(\bar{O}_x)$ is defined at every point of O_x. Thus if T is the set of points of $\bar{O}_x$ where $\bar{f}$ is not defined then $T \subseteq C_x$. Since $\operatorname{codim} C_x \geqq 2$ one also must have $\operatorname{codim} T \geqq 2$. But now for a normal affine variety Y one knows (see [3], Proposition 2, p. 166 and 10, p. 134. Also Corollary, p. 135), that if $g \in Q(Y)$ then either $g \in R(Y)$ or the set of points where g is not defined has codimension 1. Since $\bar{O}_x$ is assumed to be normal it follows that the first alternative must hold for $\bar{f}$. That is, $\bar{f}$ is everywhere defined on $\bar{O}_x$. But then $\bar{f}$, as a function on $\bar{O}_x$, is the restriction of a polynomial on X to $\bar{O}_x$. (See (2.2.2).) But then this is certainly true of f so that $f \in S(O_x)$. Q.E.D.

Remark 8. Proposition 9 is stronger than the criterion $O_x = \bar{O}_x$ for insuring $S(O_x) = R(O_x)$. In fact if $O_x = \bar{O}_x$ (in which case we may take (2) to be trivially satisfied) then C_x is empty and $\bar{O}_x$ is non-singular. But

since non-singularity implies normality the conditions of Proposition 9 are satisfied in case O_x is closed.

The proof that $\bar{O}_x$ is normal for the example of Remark 2 where $(x,x)=0$, $x\neq 0$, and $d\geqq 3$ follows from a result of Seidenberg (see § 5.1).

3. The orbit structure for the adjoint representation. 1. Let $\mathfrak{g}$ be a complex reductive Lie algebra of dimension n. Then $\mathfrak{g}$ is a Lie algebra direct sum

$$\mathfrak{g}=\mathfrak{z}+[\mathfrak{g},\mathfrak{g}] \tag{3.1.1}$$

where $\mathfrak{z}$ is the center of $\mathfrak{g}$. The commutator $[\mathfrak{g},\mathfrak{g}]$ is, as one knows, the maximal semi-simple ideal in $\mathfrak{g}$.

A subalgebra $\mathfrak{a}\subseteq\mathfrak{g}$ is said to be reductive in $\mathfrak{g}$ if the adjoint representation of $\mathfrak{a}$ on $\mathfrak{g}$ is completely reducible. Such a subalgebra is necessarily reductive (in itself).

Let $\mathfrak{g}^x$, for any $x\in\mathfrak{g}$, denote the centralizer of x. An element $x\in\mathfrak{g}$ is called semi-simple if $\operatorname{ad}x$ is diagonalizable. One knows that $\mathfrak{g}^x$ is reductive in $\mathfrak{g}$ for any semi-simple element $x\in\mathfrak{g}$ (see e.g. Theorem 7 in [11]).

An element $x\in\mathfrak{g}$ is called nilpotent in case (1) $x\in[\mathfrak{g},\mathfrak{g}]$ and (2) $\operatorname{ad}x$ is a nilpotent endomorphism.

Remark 9. If $x\in\mathfrak{a}\subseteq\mathfrak{g}$ where $\mathfrak{a}$ is reductive in $\mathfrak{g}$ then x is semi-simple (resp. nilpotent) with respect to $\mathfrak{a}$ if and only if it is semi-simple (resp. nilpotent) with respect to $\mathfrak{g}$. The proof of these statements are immediate consequences of the representation theory of reductive Lie algebras.

Now one knows that the most general element $x\in\mathfrak{g}$ may be uniquely written

$$x=y+z \tag{3.1.2}$$

where y is semi-simple, z is nilpotent and $[y,z]=0$. We will speak of y and z, respectively, as the semi-simple and nilpotent components of x. See [11], Theorem 6.

Remark 10. If $x\in\mathfrak{a}\subseteq\mathfrak{g}$ where $\mathfrak{a}$ is a subalgebra reductive in $\mathfrak{g}$ then by Remark 9 the decomposition (3.1.2) formed in $\mathfrak{g}$ is the same as the decomposition (3.1.2) formed in $\mathfrak{a}$.

In particular given the decomposition (3.1.2) one should observe that z is not only nilpotent in $\mathfrak{g}$ but also in the "reductive in $\mathfrak{g}$" subalgebra $\mathfrak{g}^y$. In particular then

$$z\in[\mathfrak{g}^y,\mathfrak{g}^y]. \tag{3.1.3}$$

Conversely if $y \in \mathfrak{g}$ is semi-simple and z is nilpotent in $\mathfrak{g}^y$ and one puts $x = y + z$ then y and z are respectively the semi-simple and nilpotent components of x.

3.2. We wish to apply the considerations of §§1 and 2 to the case where $X = \mathfrak{g}$ and $G \subseteq \operatorname{Aut} \mathfrak{g}$ is the adjoint group of $\mathfrak{g}$. Thus not only is G a connected algebraic reductive linear group but in fact G is then a semi-simple Lie group whose Lie algebra is isomorphic to $[\mathfrak{g}, \mathfrak{g}]$.

In this case we observe that the orbit O_x defined by any $x \in \mathfrak{g}$ is just the set of elements of $\mathfrak{g}$ that are conjugate to x.

If $\mathfrak{a} \subseteq \mathfrak{g}$ is any subalgebra then under the adjoint representation $\mathfrak{a}$ corresponds to a Lie subgroup $A \subseteq G$. Indeed A is the group generated by all $\exp \operatorname{ad} x$ where x ranges over $\mathfrak{a}$. In this way $\mathfrak{g}^x$ clearly corresponds to the identity component of the algebraic subgroup G^x.

We recall that an element $x \in \mathfrak{g}$ is semi-simple if and only if x may be embedded in a Cartan subalgebra (C. S.) of $\mathfrak{g}$. Equivalently $x \in \mathfrak{g}$ is semi-simple if and only if $\mathfrak{g}^x$ contains a C. S. of $\mathfrak{g}$.

The following lemma is known. We will prove it for completeness and also because, as noted in Remark 11 below, the proof may be used to give a more general result.

LEMMA 5. *Assume $x \in \mathfrak{g}$ is semi-simple. Then* (1) *G^x is connected and* (2) *O_x is closed in $\mathfrak{g}$.*

Proof. We first show G^x is connected. Let $b \in G^x$. Then by Theorem 2, p. 108, in [6], one knows that b may be uniquely written

$$b = a \exp \operatorname{ad} y \tag{3.2.1}$$

where $a \in G$ is diagonalizable and $y \in \mathfrak{g}$ is nilpotent and $a(y) = y$. Put $c_t = \exp t \operatorname{ad} x$. Then $b = c_t b c_t^{-1} = (c_t a v_t^{-1}) \exp \operatorname{ad} c_t(y)$. By the uniqueness of the decomposition (3.2.1) it follows that $a = c_t a c_t^{-1}$ and $c_t(y) = y$. Hence $a \in G^x$ and $y \in \mathfrak{g}^x$. But then b is "connected" to a in G^x by means of the curve $a \exp s \operatorname{ad} y$, $s \in \boldsymbol{R}$. Thus we may assume that b is diagonalizable. But now by Theorem 10, p. 117 in [6], if $\mathfrak{g}^b$ is the Lie subalgebra of all y such that $b(y) = y$ then $\mathfrak{g}^b$ contains a C. S. $\mathfrak{h}$ of $\mathfrak{g}$ and if $\mathfrak{h}$ is any C. S. in $\mathfrak{g}^b$ then $b = \exp \operatorname{ad} z$ for some $z \in \mathfrak{h}$.

But now $x \in \mathfrak{g}^b$ and since $\operatorname{ad} x \mid \mathfrak{g}^b$ is semi-simple there exists a C. S. $\mathfrak{h}$ such that $x \in \mathfrak{h} \subseteq \mathfrak{g}^b$. But $b = \exp \operatorname{ad} z$ for some $z \in \mathfrak{h}$. However since $\mathfrak{h} \subseteq \mathfrak{g}^x$ it follows that b may be joined to the identity in G^x by a curve; indeed one uses the curve $\exp t \operatorname{ad} z$. Hence G^x is connected.

To show that O_x is closed let $\mathfrak{h}$ be a Cartan subalgebra such that $x \in \mathfrak{h}$. By the Iwasawa decomposition we may write $G = KMH_0$ where K and M are connected Lie groups which are, respectively, compact and unipotent (an endomorphism u is called unipotent if $u - 1$ is nilpotent; a group is called unipotent if all its elements are unipotent) and H_0 is an abelian Lie group corresponding to a subalgebra of $\mathfrak{h}$. Since $x \in \mathfrak{h}$ it follows then that x is fixed under H_0. Thus $O_x = KMx$. We have proved (unpublished) that any orbit of a connected unipotent Lie group is closed. Rosenlicht [14] has generalized this to the case of a field of arbitrary characteristic. Thus we may use the reference [14] to establish that Mx is a closed subset of $\mathfrak{g}$. But since O_x is obtained by applying a compact group to a closed set it follows easily that O_x is closed.

Remark 11. Another proof that O_x is closed if x is semi-simple follows from Theorem 4, § 3.8. In fact one sees there that O_x is closed if and only if x is semi-simple. This observation was also made in [1]. Note however the proof given above, that O_x is closed when x is semi-simple generalizes and shows that the orbit of any zero weight vector for any representation of G is closed.

As a consequence of the connectivity of G^x for x semi-simple one has

LEMMA 6. *Assume $x \in \mathfrak{g}$ is semi-simple. Then $\mathfrak{g}^x$ is stable under G^x and the restriction of G^x to $\mathfrak{g}^x$ is the adjoint group of $\mathfrak{g}^x$.*

Proof. It is trivial that $\mathfrak{g}^x$ is stable under G^x. Furthermore as we have observed in the beginning of this section the identity component of G^x corresponds to $\mathfrak{g}^x$ under the adjoint representation of $\mathfrak{g}$ and hence its restriction to $\mathfrak{g}^x$ is the adjoint group of $\mathfrak{g}^x$. But G^x is connected by Lemma 5. Q. E. D.

3.3. Now for the case at hand S is just the symmetric algebra $S^*(\mathfrak{g})$ over the dual space to $\mathfrak{g}$. The well known description of the ring of invariants J given below is due to Chevalley.

If l is the rank of $\mathfrak{g}$ then J is generated by l algebraically independent homogeneous polynomials. That is, there exist homogeneous elements $u_i \in J$, $i = 1, \cdots, l$, such that if $C[Y_1, \cdots, Y_l]$ denotes the polynomial ring, over C, in l indeterminates and

$$C[Y_1, \cdots, Y_l] \to J \tag{3.3.1}$$

is the homomorphism given by $p(Y_1, \cdots, Y_l) \to p(u_1, \cdots, u_l)$ then (3.3.1) is an isomorphism. Moreover, if we write $\deg u_i = m_i + 1$ then the integers

m_i, called the exponents of $\mathfrak{g}$, are those special integers such that $\prod_{j=1}^{l}(1+t^{2m_j+1})$ is the Poincaré polynomial of $\mathfrak{g}$.

Throughout we will assume that the u_i are ordered so that

$$m_1 \leqq m_2 \leqq \cdots \leqq m_l$$

We will refer to the u_i, $i=1,2,\cdots,l$, as the primitive invariants.

Remark 12. One knows that the primitive invariants and even the l-dimensional space they span is not unique. However, in § 5.4 in connection with G-harmonic polynomials one normalizes the space they span in a natural way. See Remark 26, § 5.4.

We now define a mapping

$$u\colon \mathfrak{g} \rightarrow \boldsymbol{C}^l \tag{3.3.2}$$

by putting

$$u(x) = (u_1(x),\cdots,u_l(x)).$$

It is obvious that u is a morphism.

Now let $\boldsymbol{\mathcal{O}}$ be the set of all obits $O \subseteq \mathfrak{g}$. Since u obviously maps any orbit into a point it is clear that u induces a map

$$\eta\colon \boldsymbol{\mathcal{O}} \rightarrow \boldsymbol{C}^l$$

Now if $\mathfrak{u} \subseteq \mathfrak{g}$ is any subset stable under the action of G it is obvious that $\mathfrak{u}$ is a union of orbits. Let

$$\boldsymbol{\mathcal{O}}_{\mathfrak{u}} = \{O \in \boldsymbol{\mathcal{O}} \mid O \subseteq \mathfrak{u}\}$$

and we will let $\eta_{\mathfrak{u}}$ be the restriction of η to $\boldsymbol{\mathcal{O}}_{\mathfrak{u}}$.

Let $\mathfrak{s}$ be the set of all semi-simple elements in $\mathfrak{g}$. Obviously $\mathfrak{s}$ is stable under G so that we may consider the case where $\mathfrak{u} = \mathfrak{s}$.

Now it is easy to see that η is not one-one, that is it does not separate all orbits. One observes, however, that not only does η separate the orbits in $\mathfrak{s}$ but also that $\eta_{\mathfrak{s}}$ is a surjection. The following proposition is no doubt known. We prove it for completeness.

PROPOSITION 10. *Let $\mathfrak{s}$ be the set of all semi-simple elements in $\mathfrak{g}$. Then the map*

$$\eta_{\mathfrak{s}}\colon \boldsymbol{\mathcal{O}}_{\mathfrak{s}} \rightarrow \boldsymbol{C}^l$$

induced by u (see (3.3.2)) is a bijection.

Proposition 10 permits us to parameterize $\boldsymbol{\mathcal{O}}_{\mathfrak{s}}$ by all complex l-tuples. In

order to prove Proposition 10 we need some further notation and Lemma 7 below.

Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ regarded as fixed once and for all. Let W be the Weyl group of $\mathfrak{g}$ regarded as operating in $\mathfrak{h}$. Let $\Delta \subseteq S^1(\mathfrak{h})$ be the set of roots and let $\Delta_+ \subseteq \Delta$ be a system of positive roots fixed once and for all.

An element $x \in \mathfrak{g}$ is called regular if $\mathfrak{g}^x$ is a Cartan subalgebra. If $x \in \mathfrak{h}$ one knows that x is regular if and only if $\langle x, \phi \rangle \neq 0$ for all $\phi \in \Delta$. Now let

$$u_{\mathfrak{h}} : \mathfrak{h} \to \boldsymbol{C}^l$$

be the restriction of u to $\mathfrak{h}$.

LEMMA 7. *The map $u_{\mathfrak{h}}$ is proper. (That is, the inverse image of any compact set is compact).*

Proof. Let $\pi : \mathfrak{g} \to \operatorname{End} V$ be a faithful completely reducible representation of $\mathfrak{g}$ and let $m = \dim V$. For any positive number k let r_k be a positive number such that for any monic polynomial $Y^m + \sum_{i=0}^{m-1} c_i Y^i = p(Y)$ in the indeterminate Y, where $c_i \in \boldsymbol{C}$, one has $|c_i| \leqq k$, $i = 0, 1, \cdots, m-1$, implies $|\lambda| \leqq r_k$ for any root λ of $p(Y)$. In fact, it suffices to take $r_k = mk + 1$.

Now let $f_i \in J$ be the invariant polynomials defined so that

$$\det(Y - \pi(x)) = Y^m + \sum_{i=0}^{m-1} f_i(x) Y^i \tag{3.3.3}$$

for any $x \in \mathfrak{g}$. Now there exist unique polynomials

$$p_i(Y_1, \cdots, Y_l) \in C[Y_1, \cdots, Y_l], \; i = 0, 1, \cdots, m-1,$$

so that $f_i = p_i(u_1, \cdots, u_l)$. Thus regarding $C[Y_1, \cdots, Y_l]$ as the polynomial ring on $\boldsymbol{C}^l$ it follows that

$$f_i(x) = p_i(u_{\mathfrak{h}}(x)) \tag{3.3.4}$$

for any $x \in \mathfrak{h}$.

Now let $E \subseteq \boldsymbol{C}^l$ be any compact set. We wish to show that $u_{\mathfrak{h}}^{-1}(E)$ is compact. Let

$$k = \sup_{\substack{\xi \in E \\ i=0,1,\cdots,m-1}} |p_i(\xi)|.$$

It follows therefore from (3.3.4) that $|f_i(x)| \leqq k$ for all $x \in u_{\mathfrak{h}}^{-1}(E)$. Hence if λ is a root of (3.3.3) it follows that $|\lambda| \leqq r_k$ for any $x \in u_{\mathfrak{h}}^{-1}(E)$.

Now if $\Delta(\pi) \subseteq S^1(\mathfrak{h})$ is the set of weights of π and we put

$$|x| = \max_{\psi \in \Delta(\pi)} |\psi(x)|$$

then since π is faithful it is clear that $|x|$ defines a norm on the space $\mathfrak{h}$. But for any $\psi \in \Delta(\pi)$, $\lambda = \psi(x)$ is a root of (3.3.3). Hence $|x| \leqq r_k$ for any $x \in u_{\mathfrak{h}}^{-1}(E)$. That is, $u_{\mathfrak{h}}^{-1}(E)$ lies in the ball $|x| \leqq r_k$ and consequently is compact. Q.E.D.

Proof of Proposition 10. Let x_i, $i = 1, 2, \cdots, l$, be a basis of $\mathfrak{h}$ and let $r = \operatorname{Card} \Delta_+$. It is then a well-known result that (the generalization from the semi-simple to the reductive case is trivial)

$$\det \partial_{x_i} u_j \,|\, \mathfrak{h} = c \prod_{\phi \in \Delta_+} \phi \tag{3.3.5}$$

where c is a non-zero scalar. See e.g. [17]. But

$$c \prod_{\phi \in \Delta_+} \phi(x) \neq 0 \text{ if and only if } x \text{ is regular.} \tag{3.3.6}$$

Hence the Jacobian (3.3.5) of $u_{\mathfrak{h}}$ does not vanish identically on $\mathfrak{h}$ and consequently the Zariski closure of $u_{\mathfrak{h}}(\mathfrak{h})$ equals C^l. But since $u_{\mathfrak{h}}$ is a morphism $u_{\mathfrak{h}}(\mathfrak{h})$ contains a Zariski open subset of its closure; that is, a Zariski open subset of C^n. But any Zariski open set is dense in the usual topology. Thus in the usual sense

$$\overline{u_{\mathfrak{h}}(\mathfrak{h})} = C^l$$

But now let $\xi \in C^l$ and let $y_j \in \mathfrak{h}$, $j = 1, 2, \cdots$, be such that $u_{\mathfrak{h}}(y_j)$ converges to ξ. Since $u_{\mathfrak{h}}$ is proper the set y_j has a cluster point y. Obviously $u_{\mathfrak{h}}(y) = \xi$ and hence $u_{\mathfrak{h}}$ is surjective. It follows therefore that $\eta_{\mathfrak{s}}$ is surjective since every element of $\mathfrak{h}$ is semi-simple.

To show that $\eta_{\mathfrak{s}}$ is injective we must show that if $x, y \in \mathfrak{s}$ and $u(x) = u(y)$ then x and y are conjugate. Since every element in $\mathfrak{s}$ is conjugate to an element in $\mathfrak{h}$ we may assume $x, y \in \mathfrak{h}$. But now by Lemma 9.2 in [13] (the extension to the reductive case in trivial) if $u_j(x) = u_j(y)$ for $j = 1, 2, \cdots, l$, then x and y are conjugate under the Weyl group and consequently are conjugate with respect to G. One explicitly uses here the well known theorem that, under the mapping $S^*(\mathfrak{g}) \to S^*(\mathfrak{h})$ induced by injection $\mathfrak{h} \to \mathfrak{g}$, J maps onto the algebra of Weyl group invariants. Q.E.D.

Since $\eta_{\mathfrak{s}}$ is a bijection we may invert it. For any $\xi \in C^l$ we will let $O^{\mathfrak{s}}(\xi)$ be the unique semi-simple orbit O such that $\eta_{\mathfrak{s}}(O) = \xi$.

3.4. We now wish to look at the orbits of maximal dimension. By

adding and subtracting the dimension of the center of $\mathfrak{g}$ it is obvious from that for any $x \in \mathfrak{g}$

$$\begin{aligned} \dim O_x &= \dim \mathfrak{g} - \dim \mathfrak{g}^x \\ &= n - \dim \mathfrak{g}^x. \end{aligned} \tag{3.4.1}$$

PROPOSITION 11. *Let $x \in \mathfrak{g}$ be arbitrary. Then $\mathfrak{g}^x$ contains an l dimensional commutative subalgebra.*

Proof. This is just Theorem 5.7 in [13]. (Using the grassmannian of all l-planes in $\mathfrak{g}$ the proof is an easy consequence of the fact that the set of regular elements is dense in $\mathfrak{g}$.) Q. E. D.

As a corollary one has

PROPOSITION 12. *Let $x \in \mathfrak{g}$. Then*

$$\dim O_x \leqq n - l \tag{3.4.2}$$

and the set of x for which equality holds in (3.4.2) *is not empty.*

Proof. The equality in (3.4.2) clearly holds for all regular elements in $\mathfrak{g}$. (It also holds for a larger collection of elements. See (3.4.3) and Theorem 2, § 3.5). The inequality (3.4.2) for all elements follows from (3.4.1) and Proposition 11. Q. E. D.

By Proposition 12 $n - l$ is the maximal dimension of any orbit. We now wish to consider the set of elements in $\mathfrak{g}$ which define orbits of this dimension. Put

$$\mathfrak{r} = \{ x \in \mathfrak{g} \mid \dim \mathfrak{g}^x = l \}. \tag{3.4.3}$$

It is obvious that $\mathfrak{r}$ is stable under G so that we may consider the subset $\mathcal{O}_\mathfrak{r} \subseteq \mathcal{O}$. In fact $\mathcal{O}_\mathfrak{r}$ is the set of orbits given by

$$\mathcal{O}_\mathfrak{r} = \{ O \in \mathcal{O} \mid \dim O = n - l \}. \tag{3.4.4}$$

The structure of $\mathfrak{r}$ and $\mathcal{O}_\mathfrak{r}$ will be known as soon as we establish certain facts about principal nilpotent elements (Theorem 1). We will use a different definition of principal nilpotent than that given in [13].

Let $\mathfrak{p}$ be the set of all nilpotent elements in $\mathfrak{g}$. An element $x \in \mathfrak{g}$ is called principal nilpotent if and only if

$$x \in \mathfrak{p} \cap \mathfrak{r}$$

that is, if and only if (1) x is nilpotent and (2) $\dim \mathfrak{g}^x = l$.

Obviously $\mathfrak{p}$ is stable under G so that we may consider the set of orbits $\boldsymbol{\mathcal{O}}_{\mathfrak{p}}$. The following was established in [13].

THEOREM 1. *There are only a finite number of elements in* $\boldsymbol{\mathcal{O}}_{\mathfrak{p}}$. *Furthermore the set* $\mathfrak{p} \cap \mathfrak{r}$ *of all principal nilpotent is not empty and is in fact a single orbit* $O \in \boldsymbol{\mathcal{O}}_{\mathfrak{p}}$. *That is, for any* $e \in \mathfrak{p} \cap \mathfrak{r}$ *one has* (1)

$$\dim O_e = n - l$$

and (2)

$$\mathfrak{p} \cap \mathfrak{r} = O_e. \tag{3.4.5}$$

Moreover, this orbit is dense in $\mathfrak{p}$. *That is,*

$$\mathfrak{p} = \bar{O}_e \tag{3.4.6}$$

for any principal nilpotent element e.

If $O \in \boldsymbol{\mathcal{O}}_{\mathfrak{p}}$ *is any orbit other than the orbit of principal nilpotent elements one has*

$$\dim O < n - l. \tag{3.4.7}$$

Proof. Theorem 1 above is a restatement of Corollaries 5.3 and 5.5 in [13]. Q.E.D.

Remark 13. It follows from (3.4.6) that the set of nilpotent elements $\mathfrak{p} \subseteq \mathfrak{g}$ is an affine variety of dimension $n - l$. With regard to all the orbits in $\boldsymbol{\mathcal{O}}_{\mathfrak{p}}$ it should perhaps be recalled that in [13] it was shown that excluding the orbit consisting of zero alone they are in a natural one-one correspondence with the conjugacy classes of all 3-dimensional simple subalgebras of $\mathfrak{g}$.

If $x \in \mathfrak{g}$ is arbitrary and $x = y + z$ is the decomposition (3.1.2) then by the uniqueness of the decomposition, clearly,

$$G^x = G^y \cap G^z \tag{3.4.8}$$

and hence

$$\mathfrak{g}^x = \mathfrak{g}^y \cap \mathfrak{g}^z. \tag{3.4.9}$$

The subset $\mathfrak{r}$ may be characterized as follows:

PROPOSITION 13. *Let* $x \in \mathfrak{g}$ *be arbitrary. Write* $x = y + z$ *where* y *and* z *are, respectively, the semi-simple and nilpotent components of* x.

Then $x \in \mathfrak{r}$ *if and only if* z *is a principal nilpotent of the reductive Lie algebra* $\mathfrak{g}^y$.

Proof. Now $\mathfrak{g}^y$ is a reductive Lie algebra of rank l, and z is a nilpotent

element of $\mathfrak{g}^y$. (Remark 10, § 3.1.), Furthermore $\mathfrak{g}^y \cap \mathfrak{g}^z$ is exactly the centralizer of z in $\mathfrak{g}^y$. Thus by definition z is a principal nilpotent element of $\mathfrak{g}^y$ if and only if $\dim \mathfrak{g}^y \cap \mathfrak{g}^z = l$. But then by (3.4.9) one has $x \in \mathfrak{r}$ if and only if z is a principal nilpotent element of $\mathfrak{g}^y$. Q.E.D.

3.5. Now let $x \in \mathfrak{g}$ and let $x = y + z$ be the decomposition (3.1.2) for x. Then as was observed in [13]

$$f(x) = f(y) \tag{3.5.1}$$

for any invariant polynomial $f \in J$. That is, $f(x)$ depends only upon the semi-simple component of x. Indeed (3.5.1) is an immediate consequence of the fact (see [13], p. 1031) that

$$y + z \text{ and } y + cz \text{ are conjugate} \tag{3.5.2}$$

for any non-zero complex number c.

We can now completely describe $\boldsymbol{O}_{\mathfrak{r}}$. Proposition 10, § 3.3, shows that the orbits of semi-simple elements are in a natural one-one correspondence with $\boldsymbol{C}^l$. We now observe (and this is more important for us) that the set $\boldsymbol{O}_{\mathfrak{r}}$ of all orbits of maximal dimension $(n - l)$ is also in a natural one-one correspondence with $\boldsymbol{C}^l$.

THEOREM 2. *The map*

$$\eta_{\mathfrak{r}} \colon \boldsymbol{O}_{\mathfrak{r}} \to \boldsymbol{C}^l$$

(given by the primitive invariant polynomials $u_1, \cdots, u_l$; see § 3.3) is a bijection.

Proof. Let $\xi \in \boldsymbol{C}^l$. Then by Proposition 10, § 3.3, there exists a semi-simple element $y \in \mathfrak{g}$ such that $u(y) = \xi$. Now let z be a principal nilpotent element in $\mathfrak{g}^y$. Then by Proposition 13, and (3.5.1) if $x = y + z$ then $x \in \mathfrak{r}$ and $f(x) = f(y)$ for all $f \in J$. Hence $u(x) = \xi$. Thus $\eta_{\mathfrak{r}}(O_x) = \xi$ so that $\eta_{\mathfrak{r}}$ is surjective.

To show that $\eta_{\mathfrak{r}}$ is injective we must show that if $x_1, x_2 \in \mathfrak{r}$ and $u(x_1) = u(x_2)$ then x_1 and x_2 are conjugate. Let $x_i = y_i + z_i$, $i = 1, 2$, be the decomposition (3.1.2) for x_i. By (3.5.1) one has $u(y_1) = u(y_2)$. But then, by Proposition 10, § 3.3, y_1 and y_2 are conjugate. Hence we may assume that $y_1 = y_2 = y$. But since $x_1, x_2 \in \mathfrak{r}$ it follows from Proposition 13 that z_1 and z_2 are principal nilpotent elements of $\mathfrak{g}^y$. But then by Theorem 1 applied to $\mathfrak{g}^y$ it follows from Lemma 6, § 3.2, that there exists $a \in G^y$ such that $az_1 = z_2$. Thus $ax_1 = x_2$. Q.E.D.

Since $\eta_{\mathfrak{r}}$ is a bijection we may invert it. For any $\xi \in C^l$ let $O^{\mathfrak{r}}(\xi) \in \mathcal{O}_{\mathfrak{r}}$ be the unique orbit O of dimension $n-l$ such that $\eta_{\mathfrak{r}}(O)=\xi$.

3.6. Let $\{e_\phi\}$, $\phi \in \Delta$, be a set of root vectors belonging to Δ. Let

$$\mathfrak{m}=\sum_{\phi \in \Delta_+}(e_\phi) \tag{3.6.1}$$

be the maximal Lie algebra of nilpotent elements defined by Δ_+. Let $\mathfrak{m}^*$ be the corresponding nilpotent Lie algebra defined by the negative roots $\Delta_- = -\Delta_+$ so that

$$\mathfrak{g}=\mathfrak{m}^* + \mathfrak{h} + \mathfrak{m} \tag{3.6.2}$$

is a linear direct sum.

Let A be the Cartan subgroup of G coressponding to $\mathfrak{h}$ and let

$$\Pi = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subseteq \Delta_+$$

be the set of simple positive roots. (Here $k=l-\dim \mathfrak{z}$). For any $\phi \in \Delta$ let $a^\phi \in C^*$ be the non-zero scalar defined by

$$a(e_\phi)=a^\phi e_\phi. \tag{3.6.3}$$

Since G is the adjoint group of $\mathfrak{g}$ one knows that the mapping

$$A \rightarrow (C^*)^k \tag{3.6.4}$$

given by $a \rightarrow (a^{\alpha_1}, \cdots, a^{\alpha_k})$ is an isomorphism.

LEMMA 8. *Let $y \in \mathfrak{g}$ be semi-simple. Then the center of G^y is connected.*

Proof. If x is conjugate to y then G^x is conjugate to G^y and hence it is enough to show the center of G^x is connected for some $x \in O_y$. Since all Cartan subalgebras are conjugate we may choose $x \in \mathfrak{h}$. Moreover, by applying an element of the Weyl group W to x, if necessary, we may assume that $\mathfrak{g}^x$ is of the form

$$\mathfrak{g}^x = \mathfrak{h} + \sum_{\phi \in \Delta_1}(e_\phi) \tag{3.6.5}$$

where Π_1 is some subset of the set of simple roots Π and $\Delta_1 \subseteq \Delta$ is the set all roots ϕ of the form

$$\phi = \sum_{\alpha \in \Pi_1} n_\alpha \alpha,$$

for integers n_α.

Let Z be the center of G^x and let $a \in Z$. It is then obvious that $av = v$ for any $v \in \mathfrak{g}^x$. Hence $a \in A$ and $a^\alpha = 1$ for all $\alpha \in \Pi_1$. Since G^x is connected (see Lemma 5, § 3.2) the converse is clearly true. Hence

$$Z = \{a \in A \mid a^\alpha = 1 \text{ for all } \alpha \in \Pi_1\}.$$

Using the isomorphism (3.6.4) it is then obvious that Z is connected.

Q. E. D.

Remark 14. The structure of G^x is not as simple as Lemma 8, § 3.6 seems to indicate. In particular even though the center of G^x is connected the subgroup of G^x corresponding to the maximal semi-simple ideal $[\mathfrak{g}^x, \mathfrak{g}^x]$ of $\mathfrak{g}^x$ may have a non-trivial discrete center. The structure of G^x is analogous to that of the general linear group $Gl(d, \boldsymbol{C})$.

Let M and B be respectively the unipotent and Borel subgroups of G corresponding to $\mathfrak{m}$ and $\mathfrak{b} = \mathfrak{h} + \mathfrak{m}$.

The orbits of maximal dimension $(n - l)$ are uniform in the following respect.

PROPOSITION 14. *For any $x \in \mathfrak{r}$ the group G^x is an abelian, connected, algebraic subgroup of G of dimension $l - \dim \mathfrak{z}$. (Recall that $\mathfrak{z}$ is the center of* $\mathfrak{g}$).

Proof. By Proposition 11, § 3.4, it is immediate that $\mathfrak{g}^x$ is a commutative Lie algebra of dimension l. Therefore to prove the proposition it is enough to show that G^x is connected.

Let $e \in \mathfrak{g}$ be given by

$$e = \sum_{\alpha \in \Pi} e_\alpha$$

Then by [13], Theorem 5.3, e is a principal nilpotent element of $\mathfrak{g}$. We first show that G^e is connected. Let $h \in G^e$. But [13], Corollary 5.6, $\mathfrak{m}$ must be stable under h (this corollary asserts that a principal nilpotent element lies in one and only one nilpotent Lie algebra of the form (3.6.1)). But now since $\mathfrak{m}$ is stable under h it follows that $h \in B$. This may be proved in the following way. According to the Bruhat decomposition of G (see [7]) we may write

$$h = bs(\sigma)g$$

where $b \in B$, $g \in M$ and $s(\sigma)$ is in the normalizer of A inducing the element $\sigma \in W$ on $\mathfrak{h}$. To prove $h \in B$ it suffices to show that σ is the identity of W. But this is obvious since h, b, g and hence $s(\sigma)$ leaves $\mathfrak{m}$ stable (only the

identity element of W leaves Δ_+ stable). Thus $a \in B$. Write $h = da$ where $a \in A$ and $d \in M$. Now since $d \in M$

$$\begin{aligned} h(e) &= da(e) \\ &= a(e) + w \end{aligned}$$

where $w \in [\mathfrak{m}, \mathfrak{m}]$. But since $h(e) = e$ and

$$a(e) = \sum_{\alpha \in \Pi} a^{\alpha} e_{\alpha}$$

(see (3.6.3)) it follows that $w = 0$ and $a(e) = e$. The latter implies that $a^{\alpha} = 1$ for all $\alpha \in \Pi$ and hence by (3.6.4) a is the identity of G. Thus $h = d \in M$. But then h may be uniquely written $h = \exp \operatorname{ad} v$ where $v \in \mathfrak{m}$. But then by well known properties of nilpotent Lie algebras one has $v \in \mathfrak{g}^e$ and hence h lies in the identity component of G^e or since h was arbitrary G^e is connected.

Since all principal nilpotents of $\mathfrak{g}$ are conjugate it follows that G^z is connected for any principal nilpotent element $z \in \mathfrak{g}$.

Now let x be an arbitrary element of $\mathfrak{r}$ and let $x = y + z$ be the decomposition (3.1.2) for x. We recall that, by Proposition 13, § 3.4, z is a principal nilpotent element of the reductive Lie algebra $\mathfrak{g}^y$. Let F be the adjoint group of $\mathfrak{g}^y$ so that applying what just proved (in the case of $\mathfrak{g}$) to $\mathfrak{g}^y$ it follows that F^z is connected. Now by Lemma 6, § 3.2, the restriction of G^y to $\mathfrak{g}^y$ induces an epimorphism

$$G^y \to F \tag{3.6.6}$$

with kernel Z equal to center of G^y. But now the full inverse image of F^z under the map (3.6.6) is just $G^y \cap G^z = G^x$ (see (3.4.8)). But since F^z is connected and Z is connected by Lemma 8, § 3.6, it follows that G^x is connected. Q.E.D.

Remark 15. Note that if $x \in \mathfrak{r}$ the abelian group G^x ranges all the way from a reductive group, (in the case where x is a regular element) to a unipotent group (in the case where x is a principal nilpotent element). In the general case G^x, for any $x \in \mathfrak{r}$, is the direct product of a abelian reductive group and an abelian unipotent group.

It seems suggestive from Lemma 5, § 3.2, and Proposition 14, that possibly G^x is always connected for any $x \in \mathfrak{g}$. This, however, is false. If $\mathfrak{g}$ is the Lie algebra of the exceptional simple Lie group G_2 and $\psi = 3\alpha + 2\beta$ is the highest root where $\Pi = \{\alpha, \beta\}$, then one can show that G^x is not connected in case $x = e_\alpha + e_\psi$. In fact one sees easily that G^x contains the non-

trivial diagonalizable element $a \in A$ where $a^{\alpha} = 1$ and $a^{\beta} = -1$ whereas on the other hand the identity component of G^{x} is unipotent. One proves the latter statement by using the first line in Table 21, p. 186 in [5] (no X_0 term) together with the argument in the proof of Theorem 3, § 3.7.

3.7. It was pointed out to us by Dixmier that the following proposition is a special case of a more general result of Kirillov (all orbits of *any* Lie group for the representation contragredient to the adjoint representation are even dimensional). The simple proof given here is due to Kirillov and is easily modified to give the more general result.

PROPOSITION 15. *For any $x \in \mathfrak{g}$ one has $\dim O_x$ is even.*

Proof. Let B be a G-invariant non-singular symmetric bilinear form on $\mathfrak{g}$. Now for any $x \in \mathfrak{g}$ let B_x be the alternating bilinear form on $\mathfrak{g}$ given by

$$B_x(y, z) = B(x, [y, z]).$$

From the invariance of B it is clear that $B_x(y, z) = 0$ for all $z \in \mathfrak{g}$ if and only if $y \in \mathfrak{g}^x$. It follows therefore that B_x defines a non-singular alternating bilinear form on $\mathfrak{g}/\mathfrak{g}^x$. But since such a bilinear form can only be carried by an even dimensional space it follows that $\dim \mathfrak{g} - \dim \mathfrak{g}^x$ is even. The proposition then follows from (3.4.1). Q.E.D.

3.8. We now consider the cone $P \subseteq \mathfrak{g}$, defined as in § 1.3 by J^+. That is, P is the set of common zeros for all the polynomials in J^+.

The following proposition was essentially proved in [13] (it was proved for the case when $\mathfrak{g}$ is semi-simple).

PROPOSITION 16. *The cone P is identical with the set $\mathfrak{p} \subseteq \mathfrak{g}$ of all nilpotent elements in $\mathfrak{g}$.*

Proof. If $x \in P$ then clearly $\operatorname{ad} x$ is nilpotent. Indeed if $f_j \in S$ is defined by $f_j(y) = \operatorname{tr}(\operatorname{ad} y)^j$ for any $y \in \mathfrak{g}$ and positive integer j one has $f_j \in J^+$ and hence $f_j(x) = 0$ for all such j implies $\operatorname{ad} x$ is nilpotent. On the other hand if $x \in P$ then also $x \in [\mathfrak{g}, \mathfrak{g}]$. In fact let $x = x_1 + x_2$ be the decomposition of x according to (3.1.1) where $x_1 \in \mathfrak{z}$ and $x_2 \in [\mathfrak{g}, \mathfrak{g}]$. To see that x_1 is zero observe that every linear functional (element of S^1) on $\mathfrak{g}$ which vanishes on $[\mathfrak{g}, \mathfrak{g}]$ lies in J^1. But then one must have $f(x) = 0$ for all such linear functionals. Hence $x \in [\mathfrak{g}, \mathfrak{g}]$ and thus $P \subseteq \mathfrak{p}$ (see the definition of nilpotent elements). But $\mathfrak{p} \subseteq P$ by (3.5.2) since $y = 0$ when x is nilpotent and $f(0) = 0$ for any $f \in J^+$. Q.E.D.

Now Theorem 1 together with Propostiion 16 above implies that (1)

P has a dense orbit (the set of principal nilpotent elements) (2) P is of dimension $n-l$ and (3) P is a finite union of orbits. On the other hand from the particular structure of J the cone P can be given by considering only the primitive invariants $u_i \in J$. That is,

$$P = \{x \in \mathfrak{g} \mid u_i(x) = 0, i = 1, 2, \cdots, l\}.$$

We now generalize Theorem 1, and thereby encompass every point of $\mathfrak{g}$, by proving a similar theorem after substituting any point in C^l for the l-tuple of zeros above.

For any $\xi \in C^l$, $\xi = (\xi_1, \cdots, \xi_l)$, let $P(\xi)$, as in Proposition 6, § 1.6, be the affine variety in $\mathfrak{g}$ given by

$$P(\xi) = \{x \in g \mid u_i(x) = \xi_i,\ i = 1, \cdots, l\}.$$

That is, with respect to the map u (see (3.3.2) one has

(3.8.1) $$P(\xi) = u^{-1}(\xi).$$

Thus

(3.8.2) $$\mathfrak{g} = \bigcup_{\xi \in C^l} P(\xi)$$

is a disjoint union.

Now recall (see end of § 3.3 and § 3.5) that $O^{\mathfrak{s}}(\xi)$ and $O^{\mathfrak{r}}(\xi)$ are, respectively, the orbit of semi-simple elements and orbit of maximal dimension corresponding to any $\xi \in C^l$ under $\eta_{\mathfrak{s}}$ and $\eta_{\mathfrak{r}}$.

Now obviously $P(\xi)$ is a union of orbits. In particular $O^{\mathfrak{s}}(\xi)$ and $O^{\mathfrak{r}}(\xi)$ are contained in $P(\xi)$. Furthermore every orbit O lies in some $P(\xi)$. Theorem 1 generalizes in the following way.

THEOREM 3. *Let $\xi \in C^l$ be arbitrary. Then $P(\xi)$ is the set of all $x \in \mathfrak{g}$ whose semi-simple component lies in $O^{\mathfrak{s}}(\xi)$.*

(3.8.3) $$P(\xi) = O^{\mathfrak{r}}(\xi) \cup \cdots \cup O^{\mathfrak{s}}(\xi)$$

is a union of a finite number of orbits. Moreover

(3.8.4) $$O^{\mathfrak{r}}(\xi) = P(\xi) \cap \mathfrak{r}$$

and $O^{\mathfrak{r}}(\xi)$ is the unique orbit of maximal dimension $(n-l)$ in $P(\xi)$ and in fact

(3.8.5) $$\dim O \leqq (n-l) - 2$$

for any other orbit in $P(\xi)$. Next (2)

(3.8.6) $$O^{\mathfrak{s}}(\xi) = P(\xi) \cap \mathfrak{s}$$

and $O^s(\xi)$ is the orbit of minimal dimension in $P(\xi)$. Finally

$$P(\xi) = \overline{O^r(\xi)} \tag{3.8.7}$$

so that the Zariski closed set $P(\xi)$ is irreducible and

$$\dim P(\xi) = n - l. \tag{3.8.8}$$

Proof. The statement that $P(\xi)$ is the set of all $x \in \mathfrak{g}$ whose semi-simple component lies in $O^s(\xi)$ is an immediate consequence of (3.5.1) and Proposition 10, § 3.3. This of course implies (3.8.6). Furthermore using (3.5.2) it follows from (3.8.5) that for any orbit $O \subseteq P(\xi)$ one has

$$O^s(\xi) \subseteq \bar{O} \tag{3.8.9}$$

so that $O^s(\xi)$ is the orbit of minimal dimension in $P(\xi)$.

Now let $y \in O^s(\xi)$ so that $\mathfrak{g}^y$ is a reductive Lie algebra. Let $\mathfrak{p}^y$ denote the set of nilpotent elements of $\mathfrak{g}^y$. Applying Theorem 1 to $\mathfrak{g}^y$ there exist k elements $z_i \in \mathfrak{p}^y$, $i = 1, \cdots, k$, such that under the adjoint group F of $\mathfrak{g}^y$ every nilpotent element in $\mathfrak{g}^y$ is conjugate to one and only one of the z_i. Furthermore Theorem 1 asserts that if N is the set of all principal nilpotent elements of $\mathfrak{g}^y$ then N is an orbit under F and $\bar{N} = \mathfrak{p}^y$.

Now put $x_i = y + z_i$, $i = 1, \cdots, k$, so that by Remark 10, § 3.1, y and z_i are, respectively, the semi-simple and nilpotent components of x_i.

We now assert that every element in $P(\xi)$ is conjugate to one and only one of the x_i. We first show that the x_i lie in different conjugate classes. Assume $ax_i = x_j$ for $a \in G$. Then by the uniqueness of the decomposition (3.1.2) one has $ay = y$ so that $a \in G^y$ and $az_i = z_j$. But, by Lemma 6, § 3.2, z_i is then conjugate to z_j under F and hence $i = j$. Now let $v \in P(\xi)$ be arbitrary and let w and e be, respectively, its semi-simple and nilpotent compotents. We show that v is conjugate to one of the x_i. By the first statement of the theorem (already proved above) there exists $a \in G$ such that $aw = y$. Hence we may assume $w = y$. But then $e \in \mathfrak{p}^y$ and hence by Lemma 6 there exists $a \in G^y$ such that $ae = z_i$ for some i so that $av = x_i$. Thus there are only a finite (k) number of orbits in $P(\xi)$.

The statement (3.8.4) and the fact that $O^r(\xi)$ is the unique orbit of maximal dimension in $P(\xi)$ is just a restatement of Theorem 2. The inequality (3.8.5) then follows by Proposition 15.

Finally since N (see above) is dense in $\mathfrak{p}^y$, by Theorem 1, each x_i is in the closure of $O^r(\xi)$ so that, clearly, one has (3.8.7) and hence also (3.8.8).

Q.E.D.

Remark 16. We can be very explicit about the number of orbits in $P(\xi)$ for any $\xi \in C^l$. Let $y \in O^s(\xi)$. By the argument above and Remark 13, § 3.4, if k is the number of orbits in $P(\xi)$ then $k-1$ is the number of conjugacy classes of three dimensional simple Lie algebras in $\mathfrak{g}^y$. But such classes have been listed by Dynkin (see [5]) for every simple Lie algebra and hence one knows k as soon as one knows the maximal semi-simple ideal $[\mathfrak{g}^y, \mathfrak{g}^y]$ in $\mathfrak{g}^y$.

We note also that (3.8.9) together with Lemma 5, § 3.2, implies that the semi-simple orbits are the only closed orbits.

COROLLARY 1. *Let $x \in \mathfrak{g}$ be arbitrary. Then $\bar{O}_x$ is a union of only a finite number of orbits. Moreover if C_x is the (Zariski closed) complement of O_x in $\bar{O}_x$ then, where* codim C_x *is, as in* § 2.2, *defined with respect to O_x, one has*

$$\operatorname{codim} C_x \geqq 2.$$

Proof. If $u(x) = \xi$ then obviously $\bar{O}_x \subseteq P(\xi)$ and since $P(\xi)$ is composed of only a finite number of orbits the same is true for $\bar{O}_x$. But if $O \subseteq C_x$ then certainly

$$\dim O_x - \dim O \geqq 2$$

by Proposition 15 and the fact that $\dim O \leqq \dim C_x < \dim \bar{O}_x = \dim O_x$. But then Corollary 1 follows immediately since C_x must be a finite union of varieties of the form $\bar{O}$ where $O \subseteq C_x$. Q.E.D.

As an immediate application one has

COROLLARY 2. *Let O be any orbit and let T be the set (Zariski closed) of all non-simple points of the affine variety $\bar{O}$. Then if* codim T *is defined with respect to $\bar{O}$ one has*

$$\operatorname{codim} T \geqq 2.$$

Proof. One knows that the set of all simple points of $\bar{O}$ is Zariski open in $\bar{O}$ and non-empty. It therefore meets O. Since G is transitive on O it follows that all the points of O are simple. Thus $T \subseteq C_x$ where $O = O_x$ and hence the result follows from Corollary 1.

Remark 17. It is suggestive from Corollary 2 that possibly $\bar{O}$ is a normal variety for any orbit O. This will be proved later (see § 5.1) for all orbits of maximal dimension. For such orbits it will also be seen that T is exactly the complement of O in $\bar{O}$.

4. The transversal l-plane $\mathfrak{v}$. 1. Now we recall that S, the ring of polynomials on $\mathfrak{g}$ and S_*, the symmetric algebra over $\mathfrak{g}$ (the ring of differen-

tial operators with constant coefficients on $\mathfrak{g}$) are G-modules and are paired by (1.1.1). By taking the differential (via the adjoint representation) they become $\mathfrak{g}$-modules and by (1.1.4) one has

$$\langle x\cdot\partial, f\rangle + \langle\partial, x\cdot f\rangle = 0 \tag{4.1.1}$$

for all $x \in \mathfrak{g}$, $\partial \in S_*$ and $f \in S$. Furthermore any $x \in \mathfrak{g}$ operates as a derivation of degree 0 of S and S_* and hence, by (4.1.1), its action is completely determined by its restriction to S_1. But the latter is given by

$$x\cdot\partial_y = \partial_{[x,y]} \tag{4.1.2}$$

for any $y \in \mathfrak{g}$.

Note that if $\partial \in S_*$ is of the form $\partial = x\cdot\partial_1$ where $x \in \mathfrak{g}$ and $\partial_1 \in S_*$ then by (4.1.1)

$$\langle\partial, f\rangle = 0 \text{ for all } f \in J. \tag{4.1.3}$$

This criterion for an element $\partial \in S_*$ to be orthogonal to J is especially convenient to use when x equals a certain element $x_0 \in \mathfrak{h}$, now to be defined.

Recall that $\Pi \subseteq \Delta_+$ is the set of simple positive roots. We now put x_0 equal to the unique element in $\mathfrak{h} \cap [\mathfrak{g}, \mathfrak{g}]$ such that (see [13], § 5.2)

$$\langle x_0, \alpha\rangle = 1 \text{ for all } \alpha \in \Pi.$$

If $\phi \in \Delta$ is arbitrary and the order $o(\phi)$ of ϕ is the integer defined by

$$o(\phi) = \sum_{\alpha\in\Pi} n_\alpha(\phi) \tag{4.1.4}$$

where

$$\phi = \sum_{\alpha\in\Pi} n_\alpha(\phi)\alpha \tag{4.1.5}$$

then clearly

$$\langle x_0, \phi\rangle = o(\phi) \tag{4.1.6}$$

and hence

$$[x_0, e_\phi] = o(\phi)e_\phi. \tag{4.1.7}$$

As usual let $\mathbf{Z}$ denote the set of all integers. For every integer $j \in \mathbf{Z}$ let

$$S_*^{(j)} = \langle\partial \in S_* \mid x_0\cdot\partial = j\partial\rangle.$$

It is obvious that $S_*^{(j)}$ is a graded subspace of S_* and since x_0 operates as a derivation of S it follows immediately from (4.1.7) that

$$S_* = \sum_{j\in\mathbf{Z}} S_*^{(j)}$$

is a direct sum and

$$S_*^{(i)}S_*^{(j)} \subseteq S_*^{(i+j)}. \tag{4.1.8}$$

Similarly let $\mathfrak{g}^{(j)}$ be the eigenspace of $\operatorname{ad} x_0$ for the eigenvalue j so that $\mathfrak{g}$ is a direct sum of the $\mathfrak{g}^{(j)}$. Since $\operatorname{ad} x_0$ is a derivation of $\mathfrak{g}$ clearly

$$[\mathfrak{g}^{(i)}, \mathfrak{g}^{(j)}] \subseteq \mathfrak{g}^{(i+j)}. \tag{4.1.9}$$

The decomposition (3.6.2) is related to x_0 in the following way.

LEMMA 9. *The nilpotent Lie algebras* $\mathfrak{m}$ *and* $\mathfrak{m}^*$ *may be expressed in terms of the eigenspaces* $\mathfrak{g}^{(j)}$ *of* $\operatorname{ad} x_0$ *as follows:*

$$\mathfrak{m} = \sum_{j>0} \mathfrak{g}^{(j)}, \qquad \mathfrak{m}^* = \sum_{j<0} \mathfrak{g}^{(j)}.$$

Moreover $\mathfrak{h} = \mathfrak{g}^{(0)} = \mathfrak{g}^{x_0}$ (i.e. x_0 *is regular*).

Proof. Obvious from (4.1.7) and the fact $o(\phi)$ is positive for positive roots ϕ and negative for negative roots ϕ. Q.E.D.

Since $S_*^{(j)}$ is in the range of the action of x_0 whenever $j \neq 0$ one has, by (4.1.3),

$$\langle \partial, f\rangle = 0 \text{ if } f \in J,\ \partial \in S_*^{(j)} \text{ where } j \neq 0. \tag{4.1.10}$$

In the obvious way the symmetric algebra $S_*(\mathfrak{u})$ over any subspace $\mathfrak{u} \subseteq \mathfrak{g}$ may be regarded as a subalgebra of S_*.

Let $\mathfrak{b}$ be the maximal solvable Lie subalgebra of $\mathfrak{g}$ given by the direct sum

$$\mathfrak{b} = \mathfrak{h} + \mathfrak{m} \tag{4.1.11}$$

(resp. put $\mathfrak{b}^* = \mathfrak{m}^* + \mathfrak{h}$).

One knows that if $\mathfrak{g} = \mathfrak{gl}(d, \boldsymbol{C})$ then $f(x)$ depends only on the diagonal entries of x in case x is a triangular matrix and $f \in J$. More generally one has

PROPOSITION 17. *Let* $x \in \mathfrak{b}^*$ *so that* $x = y + v$ *where* $y \in \mathfrak{h}$ *and* $v \in \mathfrak{m}^*$. *Then for any* $f \in J$ *one has* $f(x) = f(y)$. *In particular*

$$u(x) = u(y)$$

where u *is the map* (3.3.2).

Proof. Since J is graded we may assume $f \in J^k$. Then, by (1.1.3), $k!f(x) = \langle (\partial_x)^k, f\rangle = \langle (\partial_y + \partial_v)^k, f\rangle = \langle (\partial_y)^k + \partial, f\rangle = k!f(y) + \langle \partial, f\rangle$ where,

by binomial expansion, $\partial \in \mathfrak{m}^* \cdot S_*(\mathfrak{b}^*)$. But now by Lemma 9, § 4.1 and (4.1.8) it follows that

$$\mathfrak{m}^* \cdot S_*(\mathfrak{b}^*) \subseteq \sum_{j<0} S_*^{(j)}$$

and hence $\langle \partial, f \rangle = 0$ by (4.1.10). Thus $f(x) = f(y)$. Q.E.D.

4.2. The following simple characterization of the principal nilpotent elements in $\mathfrak{m}$ was given in [13].

THEOREM 4. *Let $e \in \mathfrak{m}$. Write*

$$e = \sum_{\phi \in \Delta_+} c_\phi e_\phi$$

then e is principal nilpotent if and only if $c_\alpha \neq 0$ for every simple root $\alpha \in \Pi$.

Proof. This is just Theorem 5.3 in [13]. Q.E.D.

Now for every simple $\alpha \in \Pi$ let c'_α be an arbitrary *non-zero* complex number (normalized in § 4.4). We isolate a particular principal nilpotent element (by Theorem 5, after interchanging the roles of Δ_+ and Δ_-) e_- by putting

$$e_- = \sum_{\alpha \in \Pi} c'_\alpha e_{-\alpha}. \tag{4.2.1}$$

The following lemma gives a very simple method for constructing elements in $\mathfrak{r}$ (in fact by Lemma 11 and Proposition 10, § 3.3, at least a representative for every orbit of maximal dimension is constructed in this way).

Recall that $\mathfrak{b}$ is the maximal solvable Lie subalgebra $\mathfrak{h} + \mathfrak{m}$.

LEMMA 10. *One has the relation*

$$e_- + \mathfrak{b} \subseteq \mathfrak{r}$$

where, we recall, $\mathfrak{r}$ is the set of all $x \in \mathfrak{g}$ such that $\dim \mathfrak{g}^x = l$.

Proof. For any $j \in \mathbf{Z}$ put

$$\mathfrak{a}^{(j)} = \mathfrak{g}^{e_-} \cap \mathfrak{g}^{(j)}.$$

Then since $e_- \in \mathfrak{g}^{(-1)}$ it is clear from (4.1.9) that

$$\operatorname{ad} e_-(\mathfrak{g}^{(j)}) \subseteq \mathfrak{g}^{(j-1)} \tag{4.2.2}$$

and hence

$$\mathfrak{g}^{e_-} = \sum_{j \in \mathbf{Z}} \mathfrak{a}^{(j)} \tag{4.2.3}$$

is a direct sum.

Since $e_- \in \mathfrak{r}$ one therefore has

$$\sum_{j \in Z} \dim \mathfrak{a}^{(j)} = l. \tag{4.2.4}$$

Now filter $\mathfrak{g}$ by putting

$$\mathfrak{g}_j = \sum_{i \geqq j} \mathfrak{g}^{(i)}. \tag{4.2.5}$$

Thus (1) $\mathfrak{g}_j \supseteq \mathfrak{g}_{j+1}$ for all j (2) $\mathfrak{g} = \mathfrak{g}_j$ for j sufficiently small and (3) $\mathfrak{g}_j = 0$ for j sufficiently big.

Now let $v \in \mathfrak{b}$ and $x = e_- + v$. But now by (4.1.9) one has

$$\operatorname{ad} v(\mathfrak{g}_j) \subseteq \mathfrak{g}_j \tag{4.2.6}$$

for all j. On the other hand since the $\mathfrak{g}_j$ induce a filtration

$$\mathfrak{g}_j^x = \mathfrak{g}^x \cap g_j$$

on $\mathfrak{g}^x$ one has

$$\sum_{j \in Z} \dim \mathfrak{g}_j^x / \mathfrak{g}_{j+1}^x = \dim \mathfrak{g}^x. \tag{4.2.7}$$

But now if

$$\mathfrak{g}_j^x / \mathfrak{g}_{j+1}^x \to \mathfrak{g}^{(j)} \tag{4.2.8}$$

is the obvious injective map induced by (4.2.5) it follows immediately from (4.2.2) and (4.2.6) that the image of (4.2.8) lies in $\mathfrak{a}^{(j)}$. Thus

$$\dim \mathfrak{g}_j^x / \mathfrak{g}_{j+1}^x \leqq \dim \mathfrak{a}^{(j)} \tag{4.2.9}$$

for any j. Comparing (4.2.4) and (4.2.7) it follows that $\dim \mathfrak{g}^x \leqq l$. But, by (3.4.2), $\dim \mathfrak{g}^x \geqq l$. Hence $\dim \mathfrak{g}^x = l$ (and hence also the equality holds in (4.2.9) for any j). Q. E. D.

LEMMA 11. *Let $y \in \mathfrak{h}$ be arbitrary. Put $x = e_- + y$. Then $x \in \mathfrak{r}$ and also $u(x) = u(y)$.*

Proof. Since $y \in \mathfrak{b}$ one has $x \in \mathfrak{r}$ by Lemma 10. On the other hand since $e_- \in \mathfrak{m}^{\div}$ it follows that $u(x) = u(y)$ by Proposition 17, § 4.1. Q. E. D.

4.3. In § 1.5 we defined the notion of a quasi-regular element $x \in X$ for the general case of a linear group G operating on a vector space X. The notion is important for us because of Proposition 5, § 1.5. The question as to which elements are quasi-regular, for the case at hand, is settled by

PROPOSITION 18. *The set $\mathfrak{r}$ (all elements $x \in \mathfrak{g}$ such that $\dim \mathfrak{g}^x = l$) is identical with the set of all quasi-regular elements in $\mathfrak{g}$.* (See § 1.5.)

Proof. In the case at hand P is the set $\mathfrak{p}$ of all nilpotent elements. If $x \in \mathfrak{g}$ is quasi-regular then by definition $P_x = \mathfrak{p}$. Hence in particular if $e \in \mathfrak{p}$ is a principal nilpotent element there exists a sequence x_j, $j = 1, 2, \cdots$, such that $x_j \in O_{c_j x}$ where $c_j \in \boldsymbol{C}^*$ and such that x_j converges to e.

Now let $k = \dim \mathfrak{g}^x$. We wish to show that $k = l$. By Proposition 12, § 3.4, it suffices to show that $k \leqq l$. Obviously $\dim \mathfrak{g}^{c_j x} = k$ so that $\dim \mathfrak{g}^{x_j} = k$ for all j. Now consider the Grassmanian (which one recalls is compact) of of all subspaces of dimension k. Let $\mathfrak{u}$ be a cluster point of the $\mathfrak{g}^{x_j}$. It then follows easily (see argument in [13], p. 1003) that $\mathfrak{u} \subseteq \mathfrak{g}^e$. But then $k \leqq l$ since $\dim \mathfrak{g}^e = l$ and $\dim \mathfrak{u} = k$. Thus $k = l$ and consequently $x \in \mathfrak{r}$.

Now conversely assume that $x \in \mathfrak{r}$. We wish to show that x is quasi-regular. Since P_x is closed and stable under G to prove $P_x = \mathfrak{p}$ it suffices by (3.4.6) to show that P_x contains a principal nilpotent element.

Now let $\xi = u(x)$. Then by Proposition 10, § 3.3, there exists $y \in \mathfrak{h}$ such that $u(y) = \xi$. Put $x_1 = e_- + y$. Then by Lemma 11, § 4.2, and Theorem 2 it follows that $x_1 \in O_x$. The same argument shows that for any $c \in \boldsymbol{C}^*$ one has that $x_c \in O_{cx}$ where $x_c = e_- + cy$. (One uses the fact that $u(x) = u(y)$ implies $u(cx) = u(cy)$; an immediate consequence of the homogeneity of the u_i.) But now, obviously, $x_c \to e_-$ as $c \to 0$. Hence $e_- \in P_x$. But since e_- is principal nilpotent this proves x is quasi-regular. Q. E. D.

4.4. Now for every simple root $\alpha \in \Pi$ let $c_\alpha \in \boldsymbol{C}^*$ be any arbitrary non-zero complex number. Let e_+ be the principal nilpotent element (see Theorem 5, § 4.2) given by

$$e_+ = \sum_{\alpha \in \Pi} c_\alpha e_\alpha. \tag{4.4.1}$$

Since e_+ is principal nilpotent one has $\dim \mathfrak{g}^{e_+} = l$.

The following description of $\mathfrak{g}^{e_+}$ proved in [13] will play a fundamental role in this paper.

THEOREM 5. *There exists a basis z_i, $i = 1, 2, \cdots, l$, of $\mathfrak{g}^{e_+}$ such that (see § 4.1)*

$$z_i \in \mathfrak{g}^{(m_i)} \tag{4.4.2}$$

where, we recall, m_i is that integer given by

$$\deg u_i = m_i + 1$$

and $u_i \in J$ is the i-th primitive invariant polynomial (see § 3.3).

In particular then

$$\mathfrak{g}^{e_+} \subseteq \mathfrak{b}. \tag{4.4.3}$$

Proof. When $\mathfrak{g}$ is simple the first statement here is just the 2nd and 3rd from the last statements of Theorem 6.7 in [13] (our z_i here is the u_i of that theorem) together with Corollary 8.7 (which shows that $k_i = m_i$) of [13].

If $\mathfrak{g}$ is semi-simple the first statement is still true since e_+ may be written

$$e_+ = \sum_i e_{+,i}$$

where the $e_{+,i}$ are principal nilpotent elements of the various simple components of $\mathfrak{g}$ and each is of the form (4.4.1) for the corresponding simple component. One then uses the fact that the exponents (the m_i) of $\mathfrak{g}$ are composed of the exponents of the various simple components of $\mathfrak{g}$.

In the general case the first statement of the theorem also follows since if $\mathfrak{z}$ is the center of $\mathfrak{g}$ then $m_i = 0$ if and only if $i \leqq \dim \mathfrak{z}$. But clearly $\mathfrak{z} = \mathfrak{g}^{e_+} \cap \mathfrak{g}^{(0)}$.

Since the m_i are non-negative integers the relation (4.4.3) follows from Lemma 9, § 4.1. Q.E.D.

Remark 18. Subject only to the conditions of Theorem 5, § 4.4, it is obvious that the basis of primitive polynomial invariants u_i does not uniquely determine the basis z_i of $\mathfrak{g}^{e_+}$. However with the further relations uncovered in § 4.6 we wish to note that the u_i do uniquely determine a basis z_i of $\mathfrak{g}^{e_+}$.

Now one knows the elements $x_\alpha \in \mathfrak{h}$, $\alpha \in \Pi$, given by $x_\alpha = [e_\alpha, e_{-\alpha}]$, form a basis of $\mathfrak{h} \cap [\mathfrak{g}, \mathfrak{g}]$. Hence we may write

$$x_0 = \sum_{\alpha \in \Pi} r_\alpha x_\alpha$$

where x_0 is defined as in § 4.1.

Now define

$$e'_- = \sum_{\alpha \in \Pi} r_\alpha / c_\alpha \, e_{-\alpha}$$

where the c_α define the principal nilpotent element e_+ (see (4.4.1)). Then as observed in [13], § 5.2, the elements e'_-, x_0 and e_+ form a basis of a principal three dimensional simple subalgebra $\mathfrak{a}_0$ of $\mathfrak{g}$. Furthermore this basis satisfies the commutation relations of an S-triple (see [13], p. 996). It is obvious then that e'_- is conjugate to e_+ and hence e'_- is a principal nilpotent element of $\mathfrak{g}$. By Theorem 4, §4.2, one must therefore have that $r_\alpha \neq 0$ for any $\alpha \in \Pi$ and hence we may normalize the c'_α of (4.2.1) by putting

$$c'_\alpha = r_\alpha / c_\alpha$$

so that $e_- = e'_-$.

Now if V is any finite dimensional irreducible module for the three

dimensional simple Lie algebra $\mathfrak{a}_0$ with respect to a representation π then one knows that V is a direct sum of $\operatorname{Ker}\pi(e_+)$ and $\operatorname{Im}\pi(e_-)$. Since any finite dimensional $\mathfrak{a}_0$-module is completely reducible the same must be true without the assumption of irreducibility. It follows therefore from (4.2.2) and (4.4.3) that by restriction to $\mathfrak{b}$ we have proved.

LEMMA 12. *Let* $\mathfrak{b}$ *be the maximal solvable subalgebra of* $\mathfrak{g}$ *given by* (4.1.11). *Then*

$$\mathfrak{b} = \mathfrak{g}^{e_+} + \operatorname{ad} e_-(\mathfrak{m}) \tag{4.4.4}$$

is a direct sum.

4.5. A subset $\mathfrak{u} \subseteq \mathfrak{g}$ will be called a plane if it is of the form $\mathfrak{u} = w + \mathfrak{a}$ where $w \in \mathfrak{g}$ is an element and $\mathfrak{a}$ is a subspace. (It is called a k-plane if $\dim \mathfrak{a} = k$.)

It is obvious that a k-plane $\mathfrak{u}$ is a k-dimensional affine subvariety of $\mathfrak{g}$. Furthermore $S(\mathfrak{u})$, the restriction of S to $\mathfrak{u}$, is the affine algebra of $\mathfrak{u}$ and in fact if, as above, $\mathfrak{u} = w + \mathfrak{a}$ then writing an arbitrary $x \in \mathfrak{u}$ in the form $x = w + \sum_{i=1}^{k} r_i(x) y_i$ where y_i is a basis of $\mathfrak{a}$ it is clear that the r_i are in $S(\mathfrak{u})$ and define a coordinate system on $\mathfrak{u}$. Moreover one obviously has

$$S(\mathfrak{u}) = \boldsymbol{C}[r_1, \cdots, r_k]. \tag{4.5.1}$$

Note also that if $f \in S$ and $g = f \mid \mathfrak{u}$ then

$$\partial g / \partial r_i = \partial_{y_i} f \mid \mathfrak{u}. \tag{4.5.2}$$

Now let

$$u_{\mathfrak{u}} \colon \mathfrak{u} \to \boldsymbol{C}^l$$

be the restriction of u to $\mathfrak{u}$. The plane $\mathfrak{u}$ will be called transversal if

$$\dim u_{\mathfrak{u}}(\mathfrak{u}) = l.$$

Since $u_{\mathfrak{u}}$ is a morphism it is clear that $\mathfrak{u}$ is transversal if and only if the functions $v_i \in S(\mathfrak{u})$, where $v_i = u_i \mid \mathfrak{u}$, are algebraically independent. But this is the case if and only if the Zariski open subset $\mathfrak{u}_0$ of $\mathfrak{u}$ given by

$$\mathfrak{u}_0 = \{x \in \mathfrak{u} \mid (dv_i)_x,\ i = 1, 2, \cdots, l, \text{ are linearly independent}\}$$

is non-empty (see proof of Proposition 6, § 1.6). But by (4.5.2) $\mathfrak{u}_0$ is the set of all points x in $\mathfrak{u}$ where the $k \times l$ matrix $(\partial_{y_j} u_i)(x)$, $i = 1, \cdots, l$, $j = 1, \cdots, k$, is of rank l. Thus (to make it independent of the basis y_j of $\mathfrak{a}$)

if $d(\mathfrak{u}) \subseteq S(\mathfrak{u})$ is the space of functions on $\mathfrak{u}$ spanned by the determinants of all $l \times l$ minors of $\partial_{y_j} u_i \mid \mathfrak{u}$ (put equal to zero if $k < l$) then

(4.5.3) $\qquad$ $\mathfrak{u}$ is transversal if and only if $\dim d(\mathfrak{u}) \geqq 1$

and

(4.5.4) $$\mathfrak{u}_0 = \{x \in \mathfrak{u} \mid g(x) \neq 0 \text{ for some } g \in d(\mathfrak{u})\}.$$

Now if $\mathfrak{u}$ is an l-plane then obviously $\dim \mathfrak{u}$ is either 1 or 0 (according as to whether $\mathfrak{u}$ is transversal or not). In case $\mathfrak{u}$ is a Cartan subalgebra, say $\mathfrak{h}$, then we have already observed that $\mathfrak{u}$ is transversal. In fact

(4.5.5) $$d(\mathfrak{h}) = (\prod_{\phi \in \Delta_+} \phi)$$

and $\mathfrak{h}_0 = \mathfrak{u}_0$ is the set of all $x \in \mathfrak{h}$ which are regular in $\mathfrak{g}$. (See (3.3.5).)

Although $\mathfrak{h}$ is transversal it is not suitable for our needs, mainly because $\mathfrak{h}_0 \neq \mathfrak{h}$.

On the other hand if we put $\mathfrak{v}$ equal to the l-plane given by

(4.5.6) $$\mathfrak{v} = e_- + \mathfrak{g}^{e_+}$$

then not only is $\mathfrak{v}$ transversal but it will also be shown (1) that $d(\mathfrak{v}) = C$ so that $\mathfrak{v}_0 = \mathfrak{v}$. Moreover it will be seen that every element of $\mathfrak{v}$ lies on an orbit of maximal dimension and every such orbit meets $\mathfrak{v}$ in one and only one point.

Remark 19. In a sense $\mathfrak{v}$ is to $\mathfrak{r}$ as $\mathfrak{h}$ is to $\mathfrak{s}$, the set of all semi-simple elements in $\mathfrak{g}$. However $\mathfrak{v}$ has the advantage in that there are no "Weyl group ambiguities" with regard to conjugation. Furthermore the restriction of J to $\mathfrak{h}$ induces only a monomorphism of J into $S(\mathfrak{h})$ whereas (by Theorem 8, § 4.7) the restriction of J to $\mathfrak{v}$ induces an isomorphism of J onto $S(\mathfrak{v})$.

Remark 19′ (added in proof). If $\mathfrak{a}$ is *any* linear complement of $\operatorname{ad} e_-(\mathfrak{m})$ in $\mathfrak{b}$ which is stable under $\operatorname{ad} x_0$ it is clear that $\mathfrak{a}$ may be substituted for $\mathfrak{g}^{e_+}$ in Theorem 5. We now wish to observe that if $\mathfrak{a}$ is substituted for $\mathfrak{g}^{e_+}$ in the definition of $\mathfrak{v}$ above (see (4.5.6)) then all the results to be proved henceforth about $\mathfrak{v}$ will still hold true. That is, the only properties of $\mathfrak{g}^{e_+}$ needed are Theorem 5 and (4.4.4). In this generality the results contain Theorem 0.10 of the Introduction. In particular they apply to the special case of the plane of companion matrices. (See Remark following Theorem 0.10.)

In order to show first that $\mathfrak{v}$ is transversal the following obvious fact will be useful. Assume that $\mathfrak{u}$ and $\mathfrak{w}$ are planes and that $\psi \colon \mathfrak{u} \to \mathfrak{w}$ is a morphism defined so that

(4.5.7) $$u_{\mathfrak{w}} \circ \psi = u_{\mathfrak{u}}.$$

Then clearly (since $u_{\mathfrak{u}}(\mathfrak{u}) \subseteq u_{\mathfrak{w}}(\mathfrak{w})$)

(4.5.8) $\qquad$ $\mathfrak{u}$ is transversal implies $\mathfrak{w}$ is transversal.

The following proposition asserts among other things that every element of the l-plane $e_- + \mathfrak{h}$ is conjugate to an element in the l-plane $\mathfrak{v}$.

PROPOSITION 19. *For each element $x \in e_- + \mathfrak{h}$ there exists a unique element $a_x \in M$ such that $a_x(x) \in \mathfrak{v}$. Furthermore the map*

(4.5.9) $$e_- + \mathfrak{h} \rightarrow M$$

given by $x \rightarrow a_x$ is a morphism.

Proof. Let e_i, $i = 1, 2, \cdots, m$, be a basis of $\mathfrak{m}$ so that $e_i \in \mathfrak{g}^{(j)}$ for some $j > 0$. (See Lemma 9, § 4.1.) In fact let $r(i)$ be that positive number such that $e_i \in \mathfrak{g}^{(r(i))}$. We may then order the basis e_i so that $r(i) \leqq r(i+1)$ for all i.

Now let $w_i = [e_i, e_-]$. Then since $\mathfrak{g}^{e_-} \subseteq \mathfrak{b}^*$ (by (4.4.3) after interchanging Δ_+ and Δ_-), so that $\mathfrak{g}^{e_-} \cap \mathfrak{m} = (0)$, it follows that

$$w_i \in \mathfrak{g}^{(r(i)-1)}$$

and the w_i form a basis of ad $e_-(\mathfrak{m})$.

Obviously one has

(4.5.10) $$M \times (e_- + \mathfrak{h}) \rightarrow e_- + \mathfrak{b}$$

for the map $(a, x) \rightarrow ax$.

Now for any $v \in e_- + \mathfrak{b}$ let $c_i(v)$, $i = 1, 2, \cdots, m$, be the scalar defined so that if v is uniquely written $v = e_- + v_1 + v_2$ where $v_1 \in \mathfrak{g}^{e_+}$ and $v_2 \in \operatorname{ad} e_-(\mathfrak{m})$, according to the decomposition (4.4.4), then

$$v_2 = \sum_{i=1}^{m} c_i(v) w_i.$$

We now make the following inductive assumption about a positive integer k. There exists k functions $g_i \in S(e_- + \mathfrak{h})$, $i = 1, 2, \cdots, k$, such that if $z \in \mathfrak{m}$ where

$$z = \sum_{i=1}^{m} b_i e_i.$$

Then one has for any $x \in e_- + \mathfrak{h}$,

$$c_j(\exp \operatorname{ad} z(x)) = 0 \text{ for all } j \leqq k$$

if and only if

$$b_i = g_i(x) \tag{4.5.11}$$

for all $i \leq k$.

We now show that the assumption holds for $k+1$. Indeed if we compute $c_{k+1} = c_{k+1}(\exp \operatorname{ad} z(x))$ where z satisfies (4.5.11) it is straightforward, using (4.1.9), to see that

$$c_{k+1} = b_{k+1} + f_0(g_1(x), \cdots, g_k(x)) + \sum_{i=1}^{l} f_i(g_1(x), \cdots, g_k(x)) r_i(x) \tag{4.5.12}$$

where f_i, $i = 0, 1, \cdots, l$, are polynomials in k variables and $r_i \in S(e_- + \mathfrak{h})$ is the same as in (4.5.1) with $\mathfrak{u} = e_- + \mathfrak{h}$.

Now consider the equation $c_{k+1} = 0$. Since (4.5.12) is linear in b_{k+1} we can obviously uniquely solve for b_{k+1} obtaining $b_{k+1} = g_{k+1}(x)$ where $g_{k+1} \in S(e_- + \mathfrak{h})$. Thus the induction assumptions hold for $k+1$. On the other hand (4.5.12) is also valid for $k = 0$ provided $f_0 = 0$ and f_i are constants for $i = 1, 2, \cdots, l$. Thus, similarly, the induction assumption holds for $k = 1$.

We have thus proved inductively that given $x \in e_- + \mathfrak{h}$ there exists a unique element $z = \sum_{i=1}^{m} g_i(x) e_i$ in $\mathfrak{m}$ such that $c_i(\exp \operatorname{ad} z(x)) = 0$ for $i = 1, 2, \cdots, m$. That is, such that $\exp \operatorname{ad} z(x) \in \mathfrak{v} = e_- + \mathfrak{g}^{e_+}$ and that furthermore $g_i \in S(e_- + \mathfrak{h})$. But if $a_x = \exp \operatorname{ad} z$ this proves the lemma since one knows the map $\mathfrak{m} \to M$ given by $z \to \exp \operatorname{ad} z$ is an algebraic isomorphism.

Q. E. D.

PROPOSITION 20. *For every $x \in e_- + \mathfrak{h}$ let $a_x \in M$ be defined as in Proposition 19. If now*

$$\rho\colon e_- + \mathfrak{h} \to \mathfrak{v}$$

is the map given by $x \to a_x(x)$ then ρ is a morphism.

Proof. Obviously the map (4.5.10) is rational and everywhere defined. But by Proposition 19 so is the map

$$e_- + \mathfrak{h} \to M \times (e_- + \mathfrak{h}) \tag{4.5.13}$$

given by $x \to (a_x, x)$. One obtains the proposition by composing (4.5.10) with (4.5.13).

We can now prove

LEMMA 13. *The l-plane $\mathfrak{v} = e_- + \mathfrak{g}^{e_+}$ is transversal.*

Proof. Let ρ be as in Proposition 20. Since $\rho(x) \in O_x$ for any $x \in e_- + \mathfrak{h}$ it is obvious that

$$u_{\mathfrak{v}} \circ \rho = u_{\mathfrak{u}}$$

where $\mathfrak{u} = e_- + \mathfrak{h}$. Thus, by (4.5.8), to prove $\mathfrak{v}$ is transversal it suffices to show that $\mathfrak{u}$ is transversal. But if

$$\tau : \mathfrak{h} \to \mathfrak{u}$$

is the map by $\tau(y) = e_- + y$ then of course τ is a morphism. On the other hand by Lemma 11, § 4.2,

$$u_{\mathfrak{u}} \circ \tau = u_{\mathfrak{h}}.$$

Thus by (4.5.8) to prove $\mathfrak{u}$ is transversal it suffices to show that $\mathfrak{h}$ is transversal. But $\mathfrak{h}$ is transversal by (4.5.5). Q. E. D.

4.6. Now let z_i, $i = 1, 2, \cdots, l$, be the basis of $\mathfrak{g}^{e_+}$ given by Theorem 5, § 4.4. We recall that $z_i \in \mathfrak{g}^{(m_i)}$. On the other hand one has $\deg u_i = m_i + 1$. Hence if we put

$$(4.6.1) \qquad g_i = \frac{(\partial_{e_-})^{m_i}}{m_i!} u_i$$

then $g_i \in S^1$. That is, g_i is just a linear functional on $\mathfrak{g}$.

LEMMA 14. *Let $1 \leqq i, j \leqq l$. Then*

$$g_i(z_j) = 0$$

whenever $m_i \neq m_j$.

Proof. Since g_i is a linear function on $\mathfrak{g}$ one has, by (1.1.2) and (1.1.3),

$$(4.6.2) \qquad m_i! g_i(z) = \langle (\partial_{e_-})^{m_i} \partial_z, u_i \rangle$$

for any $z \in \mathfrak{g}$. But since $e_- \in \mathfrak{g}^{(-1)}$ one has $(\partial_{e_-})^{m_i} \in S_*^{(-m_i)}$ by (4.1.8) and, if $z \in \mathfrak{g}^{(k)}$, then also $(\partial_{e_-})^{m_i} \partial_z \in S_*^{(k-m_i)}$. But if $k \neq m_i$, that is, if $k - m_i \neq 0$ then $g_i(z) = 0$ by (4.1.10) and (4.6.2). In particular $g_i(z_j) = 0$ if $m_j \neq m_i$. Q. E. D.

The following lemma is crucial. Recall that $\mathfrak{b}$ is the maximal solvable algebra given by (4.1.11).

LEMMA 15. *Let $1 \leqq i, j \leqq l$. Then if $m_i \leqq m_j$ the function $\partial_{z_j} u_i$ reduces to a constant on $e_- + \mathfrak{b}$. In fact*

$$(4.6.3) \qquad \partial_{z_j} u_i \mid e_- + \mathfrak{b} = g_i(z_j).$$

Furthermore, if $m_i < m_j$ then the constant is zero. That is, in this case

$$\partial_{z_j} u_i \mid e_- + \mathfrak{b} = 0. \tag{4.6.4}$$

Proof. We first observe that if $x \in e_- + \mathfrak{b}$ then for any non-negative integer k one has

$$(\partial_x)^k \in \sum_{p \geqq -k} S_*^{(p)}. \tag{4.6.5}$$

Indeed this is clear from (4.1.8) and Lemma 9, § 4.1, upon writing $x = e_- + y$ where $y \in b$ and using binomial expansion. In fact from the binomial expansion it is obvious that $(\partial_{e_-})^k$ is the component of $(\partial_x)^k$ in $S_*^{(-k)}$. That is

$$(\partial_x)^k - (\partial_{e_-})^k \in \sum_{p > -k} S_*^{(p)}. \tag{4.6.6}$$

Now if $f \in J$ it follows, since $z_j \in \mathfrak{g}^{(m_j)}$, that by (1.1.2), (4.1.8) and (4.1.10)

$$\langle \partial, \partial_{z_j} f \rangle = 0 \tag{4.6.7}$$

for all $\partial \in S_*^{(p)}$ where $p \neq -m_j$; in particular for all $p > -m_j$.

But now if $k = m_i$ in (4.6.6) then the sum there is over all p where $p > -m_i$. Hence if $m_i \leqq m_j$, so that $-m_i \geqq -m_j$, the sum in (4.6.6) is over all p, where $p > -m_j$. Thus, by (4.6.7),

$$\langle (\partial_x)^{m_i}, \partial_{z_j} f \rangle = \langle (\partial_{e_-})^{m_i}, \partial_{z_j} f \rangle \tag{4.6.8}$$

for all $x \in e_- + \mathfrak{b}$ whenever $m_i \leqq m_j$.

We now assert that this implies

$$(\partial_{z_j} u_i)(x) = g_i(z_j) \tag{4.6.9}$$

for any $x \in e_- + \mathfrak{b}$ whenever $m_i \leqq m_j$. Indeed replace f by u_i and divide by $m_i!$ in (4.6.8). Recalling that $\deg \partial_{z_j} u_i = m_i$ the left side of (4.6.8) becomes the left side of (4.6.9) by (1.1.3). On the other hand by (1.1.2) the right side of (4.6.8) becomes the right side of (4.6.9) by (4.6.2). (Recall that S_* is commutative.) This proves (4.6.3).

But now if $m_i < m_j$ then the right side of (4.6.9) vanishes by Lemma 14. Hence one obtains (4.6.4). Q.E.D.

We can now show that the Jacobian matrix of functions $\partial_{z_j} u_i \mid \mathfrak{v}$ of the map $u_{\mathfrak{v}}$ takes triangular form and reduces to non-zero constants along the diagonal.

THEOREM 6. *There exists a unique basis z_j, $j=1,2,\cdots,l$, of $\mathfrak{g}^{e_+}$ such that for $i=1,2,\cdots,l$,*

$$g_i(z_j)=\delta_{ij}. \tag{4.6.10}$$

Furthermore the basis satisfies the condition of Theorem 5. That is $z_j\in\mathfrak{g}^{(m_j)}$ for all j. Furthermore

$$\partial_{z_j}u_i \mid \mathfrak{v}=\begin{cases}0 \text{ for } i<j\\ 1 \text{ for } i=j\end{cases} \tag{4.6.11}$$

so that not only is $\mathfrak{v}$ transversal but in fact

$$\det \partial_{z_j}u_i \mid \mathfrak{v}=1 \tag{4.6.12}$$

and hence (see § 4.5)

$$d(\mathfrak{v})=\boldsymbol{C}. \tag{4.6.13}$$

Proof. An integer k will be called an exponent if $k=m_i$ for some i. Let E be the set of exponents and for any $k\in E$ let $P_k\subseteq\{1,2,\cdots,l\}$ be the set of all i such that $m_i=k$. Now, for any $k\in E$ put

$$b_k=\det_{i,j\in P_k} g_i(z_j).$$

It then follows from Lemma 15 that $\det\partial_{z_j}u_i$ is a constant on $e_-+\mathfrak{b}$ and in fact

$$\det\partial_{z_j}u_i \mid e_-+\mathfrak{b}=\prod_{k\in E} b_k.$$

But since $\mathfrak{v}\subseteq e_-+\mathfrak{b}$ and since $\mathfrak{v}$ is transversal (Lemma 13) this constant can not be zero. Thus $b_k\neq 0$ for any $k\in E$. That is, the matrix $g_i(z_j)$, $i,j\in P_k$, is non-singular and this holds for any $k\in E$. It follows immediately then from Lemma 14 that a unique basis z_j of $\mathfrak{g}^{e_+}$ exists so that (4.6.10) is satisfied. It is also clear from Lemma 14 that the z_j necessarily satisfy the condition of Theorem 5. Since $\mathfrak{v}\subseteq e_-+\mathfrak{b}$ the remaining statements follow from Lemma 15. Q.E.D.

4.7. We will assume from here on that the basis z_j of $\mathfrak{g}^{e_+}$ is given by Theorem 6. Now let $s_j\in S(\mathfrak{v})$ be the coordinate functions on $\mathfrak{v}$ corresponding to the z_j. That is, s_j is such that $x=e_-+\sum s_j(x)z_j$. We have already noted that $S(\mathfrak{v})=\boldsymbol{C}[s_1,\cdots,s_l]$ (see (4.5.1)).

In notational simplicity let

$$t\colon\mathfrak{v}\to\boldsymbol{C}^l$$

(instead of $u_{\mathfrak{v}}$) denote the restriction of u to $\mathfrak{v}$. Thus for any $x\in\mathfrak{v}$

$$t(x)=(t_1(x),\cdots,t_l(x))$$

where $t_i = u_i \mid \mathfrak{v}$. It follows therefore from (4.5.2) that

$$\frac{\partial t_i}{\partial s_j} = \partial_{x_j} u_i \mid \mathfrak{v}. \tag{4.7.1}$$

Now if $\mathfrak{u}$ is an arbitrary k-plane in $\mathfrak{g}$ let

$$J \to S(\mathfrak{u}) \tag{4.7.2}$$

be the ring homomorphism obtained by restricting an invariant polynomial to $\mathfrak{u}$. Now in general one could hardly expect (4.7.2) to be an isomorphism. Indeed if (4.7.2) is an epimorphism one must have $k \geqq l$ and if (4.7.2) is a monomorphism one must have $k \leqq l$ (since the u_i are algebraically independent). Hence the possibility could only exist if $k = l$. If $\mathfrak{u}$ is a Cartan subalgebra the one knows that (4.7.2) is a monomorphism and the image is the space of Weyl group invariants. Hence in such a case (4.7.2) is an isomorphism only when $\mathfrak{g}$ is abelian. On the other hand when $\mathfrak{u} = \mathfrak{v}$ we have, in general, the following corollary of Theorem 6

THEOREM 7. *If* $\mathfrak{u} = \mathfrak{v}$ *then* (4.7.2) *is an isomorphism. Moreover the map*

$$t\colon \mathfrak{v} \to \boldsymbol{C}^l \tag{4.7.3}$$

obtained by restricting u *to* $\mathfrak{v}$ *is an algebraic isomorphism so that* $t_1, \cdots, t_l$ *define a global coordinate system on* $\mathfrak{v}$.

Furthermore the relationship between the t_i *and the linear coordinates* s_j *on* $\mathfrak{v}$ *is as follows: For* $i = 1, 2, \cdots, l$, *there exists polynomials* p_i *and* q_i *in* $i - 1$ *variables without constant term such that*

$$t_i = s_i + p_i(s_1, \cdots, s_{i-1}) \tag{4.7.4}$$

and

$$s_i = t_i + q_i(t_1, \cdots, t_{i-1}) \tag{4.7.5}$$

Proof. To prove the theorem observe that it suffices only to prove (4.7.4). Indeed using (4.7.4) we can solve for s_i obtaining (4.7.5). It is then immediate that t is one-one, onto and is in fact a biregular birational map. Since the t_i generate the image of J in $S(\mathfrak{v})$ it is then also obvious that (4.7.2) is an isomorphism.

But now (4.7.4) is immediate from (4.6.11) and (4.7.1).

Finally by definition of the coordinate system s_i one has $s_i(e_-) = 0$ for all i. On the other hand $t_i(e_-) = 0$ for all i since $t(e_-) = u(e_-) = 0$ (recall that e_- is nilpotent). Thus the p_i and q_i have no constant term. Q.E.D.

Any orbit O of semi-simple elements (i.e., $O \in \mathfrak{O}_{\mathfrak{s}}$) intersects $\mathfrak{h}$ in a finite number but in general more than one point. We now find that any orbit O of maximal dimension (i.e. $O \in \mathfrak{O}_{\mathfrak{r}}$) intersects $\mathfrak{v}$ in one and only one point.

THEOREM 8. *One has* $\mathfrak{v} \subseteq \mathfrak{r}$. *Furthermore if*

$$\mathfrak{v} \to \mathfrak{O}_{\mathfrak{r}} \tag{4.7.6}$$

is the map given by $x \to O_x$ *then* (4.7.6) *is a bijection. That is, no two distinct elements of* $\mathfrak{v}$ *are conjugate and every element in* $\mathfrak{r}$ *is conjugate to one and only one element in* $\mathfrak{v}$.

Proof. Since $\mathfrak{v} \subseteq e_- + \mathfrak{b}$ one has $\mathfrak{v} \subseteq \mathfrak{r}$ by Lemma 10, § 4.2. But now if we compose (4.7.6) with the bijection $\eta_{\mathfrak{r}}$ (Theorem 2, § 3.5) we obtain the bijection t (see Theorem 7). Hence (4.7.6) must be a bijection. Q. E. D.

We can now obtain the following characterization of the set $\mathfrak{r}$.

THEOREM 9. *Let* $x \in \mathfrak{g}$. *Then* $x \in \mathfrak{r}$ *if and only if* $(du_i)_x$, $i = 1, 2, \cdots, l$, *are linearly independent.*

Proof. By (4.6.12) the matrix $(\partial_{z_j} u_i)(x)$ is of rank l for any $x \in \mathfrak{v}$. Thus $(du_i)_x$, $i = 1, \cdots, l$, are linearly independent for any $x \in \mathfrak{v}$. But then by Theorem 9 and conjugation the same is true for any $x \in \mathfrak{r}$.

Now let $x \in \mathfrak{g}$ but where $x \notin \mathfrak{r}$. We must prove that the $(du_i)_x$, $i = 1, 2, \cdots, l$, are linearly dependent. Assume first that $x \in \mathfrak{s}$ (that is, x is semi-simple). Then x is not regular so that $\mathfrak{g}^x$ contains a Cartan subalgebra as a proper subalgebra. It follows therefore that if $\mathfrak{u}$ is the center of $\mathfrak{g}^x$ and $l_x = \dim \mathfrak{u}$ one has $l_x < l$.

Furthermore it is also clear that $\mathfrak{u}$ is the set of fixed vectors for the action of G^x on $\mathfrak{g}$ (recall that G^x is connected. See Lemma 5, § 3.2). Thus there exists a non-abelian simple component $\mathfrak{g}_1$ of $\mathfrak{g}$ of rank, say l_1, such that in the notation of § 2.1

$$d_\psi{}^{G^x} < l_1$$

where $\psi \in D$ and the irreducible representation ν^ψ is equivalent to the adjoint action of G on $\mathfrak{g}_1$. (One uses here that ν^ψ is self-contragredient.) But then by Proposition 8 the multiplicity of ν^ψ in the G-module $R(G/G^x)$ or $R(O_x)$ is less than l_1. But $S(O_x) \subseteq R(O_x)$ (in fact here $S(O_x) = R(O_x)$ since O_x is closed. See (2.2.3)) so that the same is true for the G-module $S(O_x)$.

Now let $\mathfrak{g}_2$ be an ideal in $\mathfrak{g}$ complementary to $\mathfrak{g}_1$ so that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ is a direct sum. It is obvious that we may choose the primitive invariants so that for $1 \leqq i_1 < i_2 < \cdots < i_{l_1} \leqq l$ one has

$$u_{i_k}(x_1 + x_2) - u_{i_k}(x_1)$$

where $x_j \in \mathfrak{g}_j$, $j=1,2$, and $k=1,2,\cdots,l_1$. (The dependence or independence of the $(du_i)_x$ obviously does not depend upon the choice of the primitive generators u_i). That is, for any such k,

$$(4.7.7) \qquad \partial_y(u_{i_k})=0 \text{ for any } y\in\mathfrak{g}_2.$$

Now for any $u\in J$, $y\in\mathfrak{g}$ and $a\in G$ one clearly has

$$(4.7.8) \qquad a\cdot\partial_y u=\partial_{ay}u.$$

Now let y be the root vector e_ϕ where ϕ is the highest root of $\mathfrak{g}_1$. It follows then from (4.7.8) that $\partial_y u\in S^\psi$, and hence $\partial_y u \mid O_x\in S^\psi(O_x)$, and in fact, if not zero, these functions are highest weight vectors. But then since the multiplicity of ν^ψ in $S(O_x)$ is less than l_1 it follows that $\partial_y u_{i_k}\mid O_x$, $k=1,2,\cdots,l_1$, must be linearly dependent. Thus there exists scalars c_{i_k}, not all zero, such that if $u=\sum_k c_{i_k}u_{i_k}$ then $\partial_y u\mid O_x=0$. But then by (4.7.7) and (4.7.8) one has $\partial_z u\mid O_x=0$ for all $z\in\mathfrak{g}$. In particular then

$$(\partial_z u)(x)=0 \text{ for all } z\in\mathfrak{g}.$$

Thus $(du)_x=0$ and hence the $(du_i)_x$, $i=1,2,\cdots,l$, are linearly dependent.

Now assume that $x\in\mathfrak{g}$ is arbitrary where $x\notin\mathfrak{r}$. Since the set of all y such that the $(du_i)_y$ are linearly independent is an open set (see § 1.6) to prove the theorem it suffices from above to show that there exists a sequence x_k such that $x_k\to x$ where the x_k are semi-simple but not regular. In fact it suffices to show this for the case where x is nilpotent. Indeed if $x=y+z$ is the decomposition (3.1.2) for x then by Proposition 13 z is not a principal nilpotent element if $\mathfrak{g}^y$. Hence if such a sequence has been shown to exist in the nilpotent case there exists a sequence y_k of non-regular semi-simple elements in $\mathfrak{g}^y$ such that $y_k\to z$. Hence $x_k=y+y_k$ converges to x. But clearly x_k is semi-simple and non-regular in $\mathfrak{g}$.

Hence we may assume that x is a non-principal nilpotent element of $\mathfrak{g}$. By conjugation we may also assume $x\in\mathfrak{m}$ so that, by Theorem 4, § 4.2, $x=\sum_{\phi\in\Delta_+}c_\phi e_\phi$ where there exists a simple root $\alpha\in\Pi$ such that $c_\alpha=0$. Let

$$\mathfrak{n}=\sum_{\substack{\phi\in\Delta_+\\ \phi\neq\alpha}}(e_\phi)$$

so that $x\in\mathfrak{n}$. Also let $y\in\mathfrak{h}$ be such that $\langle\beta,y\rangle$ is positive for $\beta\in\Pi$ where $\beta\neq\alpha$ and $\langle\alpha,y\rangle=0$. It follows therefore that y is semi-simple and non-regular. Furthermore it is clear that $\mathfrak{n}$ is in the range of $\operatorname{ad}y$. In fact $[y,\mathfrak{n}]=\mathfrak{n}$ so that by the argument of Theorem 3.6 in [13] (See Note following this theorem) every element in $y+\mathfrak{n}$ is conjugate to y. In particular

$y+x$ is conjugate to y and hence is semi-simple and non-regular. But we may substitute y/k for y. Hence if $x_k=y/k+x$ then $x_k\to x$ where x_k is semi-simple and non-regular. Q. E. D.

4.8. Generalizing the definition of J^+, let J^ξ, for any $\xi\in \boldsymbol{C}^l$, be the (maximal) ideal in J generated by $u_i-\xi_i$, $i=1,2,\cdots,l$. Obviously then

$$J^\xi S=(u_1-\xi_1,\cdots,u_l-\xi_l)$$

Recall that $P(\xi)\subseteq \mathfrak{g}$ is the set of zeros of $J^\xi S$. We can now prove

THEOREM 10. *Let $\xi\in \boldsymbol{C}^l$ be arbitrary. Then $P(\xi)$ is a Zariski closed subvariety (of X) of dimension $n-l$ and its ideal $I(P(\xi))$ is given by*

$$I(P(\xi))=J^\xi S \tag{4.8.1}$$

so that (a) *$P(\xi)$ is a complete intersection and* (b) *$J^\xi S$ is a prime ideal. Furthermore if $P(\xi)_s$ is the set of simple points of $P(\xi)$ then*

$$P(\xi)_s=O^r(\xi) \tag{4.8.2}$$

where $O^r(\xi)$ is the unique orbit of dimension $n-r$ in $P(\xi)$. (*See Theorem* 4, § 3.8). *Moreover the set of non-simple points in $P(\xi)$ is a finite union of orbits and has a codimension of at least* 2 *in $P(\xi)$.*

Proof. By Theorem 3, § 3.8, $P(\xi)$ is a Zariski closed subvariety of dimension $n-l$ and $P(\xi)\cap \mathfrak{r}$ is the orbit $O^r(\xi)$ of dimension $n-l$ defined in § 3.5.

But if $x\in P(\xi)$ then $(du_i)_x$, $i=1,2,\cdots,l$, are linearly independent if and only if $x\in O^r(\xi)$ by Theorem 9, § 4.7. Hence by Proposition 6, § 1.6, one obtains (4.8.1) and (4.8.2). Furthermore the set of non-simple points of $P(\xi)$ is a finite union of orbits and have a codimension of at least two in $P(\xi)$ by Theorem 3, § 3.8. Q. E. D.

Remark 20. By (4.8.2) note that $P(\xi)$ is a non-singular variety if and only if $O^r(\xi)=P(\xi)$; that is, if and only if $O^r(\xi)$ is an orbit of regular elements.

Now let B be a non-singular symmetric G-invariant bilinear form on $\mathfrak{g}$. (One extends the Cartan-Killing form of $[\mathfrak{g},\mathfrak{g}]$ in an obvious way.)

Let $H\subseteq S$ be the graded space of G-harmonic polynomials defined as in § 1.4 so that $S=J^+S+H$ is a G-module direct sum.

Remark 21. By Proposition 16, § 3.8, note that the subspace $H_P\subseteq S$ (see § 1.4) is the space of polynomials spanned by all powers of $g^k\in S^k$, $k=0,1,\cdots$, and all linear functionals $g\in S^1$ corresponding to ∂_x, under the map (1.4.1), where x is an arbitrary nilpotent element of $\mathfrak{g}$.

Finally for any $\lambda \in D$ recall that l_λ is the multiplicity the zero weight for the representation ν^λ. That is, $l_\lambda = d_\lambda{}^A$ for any Cartan subgroup $A \subseteq G$. (See § 2.1.)

We can now prove

THEOREM 11. *Let* $\mathfrak{g}$ *be a complex reductive Lie algebra and let* S *be the ring of all polynomials on* $\mathfrak{g}$. *Let* G *be the adjoint group of* $\mathfrak{g}$ *and let* $J \subseteq S$ *be the subring of* G*-invariant polynomials. Then* S *is free as a* J*-module (under multiplication). Furthermore*

$$H = H_P \tag{4.8.3}$$

and if

$$J \otimes H \to S \tag{4.8.4}$$

is the map given by $u \otimes h \to uh$ *then* (4.8.4) *is a* G*-module isomorphism. Also for any* $\xi \in \boldsymbol{C}^l$ *the ideal* $J^\xi S$ *is prime in* S *and*

$$S = J^\xi S + H \tag{4.8.5}$$

is a direct sum.

Moreover H *is completely reducible as a* G*-module and for any* $\lambda \in D$ *the irreducible representation* ν^λ *of* G *occurs with multiplicity* l_λ *in* H *(so that* $H = \sum_{\lambda \in D} H^\lambda$ *is a direct sum and* $\dim H^\lambda = l_\lambda d_\lambda$ *where* d_λ *is the dimension of* ν^λ*).*

Let $\mathfrak{r}$ *be the set of all* $x \in \mathfrak{g}$ *whose corresponding orbit* O_x *has maximal dimension* $(n - l)$. *Let* $x \in \mathfrak{r}$ *and let* $S(O_x)$ *be the ring of functions on* $S(O_x)$ *obtained by restricting* S *to* O_x *and let*

$$H \to S(O_x) \tag{4.8.6}$$

be the map obtained by restricting G*-harmonic functions to* O_x. *Then* (4.8.6) *is a* G*-module isomorphism so that all* $S(O_x)$, *for* $x \in \mathfrak{r}$, *are isomorphic as* G*-modules.*

Proof. By Theorem 1, § 3.4, and Proposition 15, § 3.8, the cone P has a dense orbit and by Theorem 10 J^+S is a prime ideal (case where $\xi = 0$). One obtains (4.8.3) and (4.8.4) as a consequence of Proposition 4, § 1.4. Moreover the map (4.8.6) is an isomorphism by (1.5.2) and Proposition 18, § 4.3. Since $J^\xi S$ is prime by Theorem 10 one therefore obtains the direct sum decomposition (4.8.5) (using 3.8.7)).

Obviously H is a completely reducible G-module. To find the multiplicity of ν^λ in H one uses the isomorphism (4.8.6) and chooses x to be regular.

In such a case $R(O_x) = S(O_x)$ by (2.2.3) since $O_x = \bar{O}_x$. But the multiplicity of ν^λ in $R(O_x)$ is l_λ by Proposition 8 since G^x is a Cartan subgroup of G. Q.E.D.

Remark 22. Except for (4.8.3) note that by Proposition 2, § 1.3, one may replace H in Theorem 11 by any G-stable complement of J^+S in S.

Also we wish to note that every irreducible representation of G appears with positive multiplicity in H. That is $l_\lambda \geqq 1$ for any $\lambda \in D$. Indeed let $Z \subseteq \mathfrak{h}'$ be the discrete subgroup generated by all roots $\phi \in \Delta$. Since G is the adjoint group one knows that every weight of ν^λ can be regarded as an element of Z. On the other hand if D is identified with the subset of all $\mu \in Z$ such that $\mu(x_\phi) \geqq 0$, for all $\phi \in \Delta_+$, where $x_\phi \in \mathfrak{h}$ is the root normal corresponding to ϕ, in such a way that λ is the highest weight of ν^λ then it is known that any $\mu \in D$ is a weight of ν^λ if and only if $\lambda - \mu$ is a non-negative integral combination of positive roots. Since $\mu = 0$ always satisfies this condition it follows that $l_\lambda \geqq 1$.

Remark 22′. It has been pointed out to us by Serre, as we ourselves have also noticed, one of the conclusions of Theorem 11, namely $S = J \otimes H$, can be obtained directly from the theorem of Chevalley mentioned in Example 1 of the Introduction.

(*Added in proof.*) As used above, the primeness of J^+S implies that J^+S is the ideal of the variety P. Another application of this fact in algebraic geometry is the following theorem.

Clearly P meets the Cartan subalgebra $\mathfrak{h}$ (or any Cartan subalgebra) only at the origin and $\dim P + \dim \mathfrak{h} = \dim \mathfrak{g}$.

THEOREM 12. *Let w be the intersection multiplicity of P and $\mathfrak{h}$ at the origin. Then w is the order of the Weyl group.*

Proof. Let L be the local ring at the origin (as a point of $\mathfrak{g}$) and let I and K, respectively, be the prime ideals of L corresponding to $\mathfrak{h}$ and P. Now one knows that w is the alternating sum of the integers $\dim \operatorname{Tor}_i^L(L/I, L/K)$. But clearly

$$\dim \operatorname{Tor}_i^L(L/I, L/K) = \dim \operatorname{Tor}_i^S(S(\mathfrak{h}), S(P))$$

We now observe, however, that by Theorem 11 one has $\operatorname{Tor}_i^S(S(\mathfrak{h}), S(P)) = 0$ for $i > 0$. Indeed since S is J-free and $S(P) = S/J^+S$ there is a spectral sequence converging to $\operatorname{Tor}_i^S(S(\mathfrak{h}), S(P))$ where

$$E_{p,q}{}^2 = \operatorname{Tor}_p{}^{S(P)}(\operatorname{Tor}_q{}^J(S(\mathfrak{h}), \boldsymbol{C}), S(P))$$

(See Cartan and Eilenberg, Homological Algebra, Chapter XVI, § 6, Theorem

6.1, p. 349). On the other hand since $S(\mathfrak{h})$ is J-free (See [2]) one has $E_{p,q}{}^2=0$ unless $p=q=0$ and $E_{0,0}{}^2=S(\mathfrak{h})/J^+S(\mathfrak{h})$. This verifies the observation and also proves that

$$w=\dim S(\mathfrak{h})/J^+S(\mathfrak{h})\,. \tag{4.8.7}$$

But by [2] (and also from the cohomology theory of the generalized flag manifold) one knows that the right side of (4.8.7) is the order of the Weyl group. Q.E.D.

4.9. Now let $p(t)$ be the formal power series

$$p(t)=\sum_{k=0}^{\infty}\dim H^k t^k.$$

Since S is isomorphic to the tensor product $J\otimes H$ by Theorem 11 it follows therefore that

$$p(t)=\frac{\prod_{i=1}^{l}1-t^{m_i}}{(1-t)^n}. \tag{4.9.1}$$

Now let $\bar{\mathfrak{g}}$ denote the projective space of all one dimensional subspaces of $\mathfrak{g}$ and let

$$\mathfrak{g}-(0)\to\bar{\mathfrak{g}} \tag{4.9.2}$$

be the canonical projection map. If $\mathfrak{u}$ is a homogeneous Zariski closed subvariety of $\mathfrak{g}$ let $\bar{\mathfrak{u}}\subseteq\mathfrak{g}$ be the image of $\mathfrak{u}-(0)$ under (4.9.2) so that $\bar{\mathfrak{u}}$ is a projective variety.

In the remaining portion of this section we use the notation of FAC, [15].

Consider the projective variety $\bar{\mathfrak{p}}\subseteq\bar{\mathfrak{g}}$ defined by the cone $\mathfrak{p}$ of all nilpotent elements in $\mathfrak{g}$. The dimensional determination of the sheaf cohomology groups $H^j(\bar{\mathfrak{p}},\mathcal{O}(k))$ for all $j,k\in\boldsymbol{Z}$ is given by

THEOREM 13. *Let $k\in\boldsymbol{Z}$ be arbitrary. Then*

$$H^j(\bar{\mathfrak{p}},\mathcal{O}(k))=0 \tag{4.9.3}$$

where j is any integer other than 0 *or* $n-l-1$. *On the other hand*

$$H^0(\bar{\mathfrak{p}},\mathcal{O}(k))=H^{n-l-1}(\bar{\mathfrak{p}},\mathcal{O}(\frac{l-(2k+n)}{2})). \tag{4.9.4}$$

(Recall that $n-l$ is even.)

Furthermore if $q(t)$ is the formal power series defined by

$$q(i)=\sum_{k=-\infty}^{\infty}\dim H^0(\bar{\mathfrak{p}},\mathcal{O}(k))\,t^k$$

then $q(t)=p(t)$. *That is* $q(t)$ *may be given by*

$$(4.9.5)\qquad q(t)=\frac{\prod_{i=1}^{l}1-t^{m_i}}{(1-t)^n}.$$

Proof. By Proposition 15, § 3.8, one has $\mathfrak{p}=P$. Hence by Theorem 10, for $\xi=0$, one has that $\bar{\mathfrak{p}}$ is a complete intersection so that Proposition 5, § 78, in [15] is applicable. This yields (4.9.3) and (4.9.4) since clearly $N=l-n$. On the other hand in the notation of [15] one has $\dim H^0(\bar{\mathfrak{p}},\mathcal{O}(k))=\dim S^k(\bar{\mathfrak{p}})$. But since J^+S is the prime ideal of $\bar{\mathfrak{p}}$ by (4.8.1) and since $S=J^+S+H$ is a direct sum it follows that $\dim H^k=\dim S^k(\bar{\mathfrak{p}})$. Thus $p(t)=q(t)$ and hence one obtains (4.9.5) from (4.9.1). Q. E. D.

4.10. One obviously has an isomorphism $f\to f^*$ of J onto the ring $S(\boldsymbol{C}^l)$ of all polynomials on $\boldsymbol{C}^l$ by defining, for any $\xi\in\boldsymbol{C}^l$,

$$f^*(\xi)=p(\xi_1,\cdot\cdot\cdot,\xi_l)$$

where p is that polynomial in l variables such that $f=p(u_1,\cdot\cdot\cdot,u_l)$. We recall that the u_i are the primitive invariant polynomials.

Now let $\mathcal{U}$ be the set of all Zariski closed subvarieties of $\boldsymbol{C}^l$. We may use $\mathcal{U}$ to index all the prime ideals in J by defining $J^U\subseteq J$ for $U\in\mathcal{U}$ to be the prime ideal consisting of all $f\in J$ such that f^* vanishes on U.

If $I\subseteq S$ is any prime ideal in S let $\mathfrak{u}(I)\subseteq\mathfrak{g}$ be the corresponding Zariski closed subvariety of $\mathfrak{g}$ of all points in $\mathfrak{g}$ at which I vanishes. It is of course obvious that I is stable under G if and only if $\mathfrak{u}(I)$ is stable under the action of G on $\mathfrak{g}$; that is, if and only if $\mathfrak{u}$ is a union of orbits. It is clear of course that if I is generated by invariant polynomials then I is G-stable. However this is not a necessary condition. The question arises: how does one characterize all those G-stable Zariski closed subvarieties $\mathfrak{u}$ of $\mathfrak{g}$ whose prime ideal $I(\mathfrak{u})$ is generated by invariant polynomials? The following theorem asserts that a necessary and sufficient condition is $\mathfrak{u}\cap\mathfrak{r}$ should not be empty. Note that since $\mathfrak{r}$ is a Zariski open subset of $\mathfrak{g}$ (obvious from its definition. Also see the proof of Proposition 6, § 1.6) then $\mathfrak{u}\cap\mathfrak{r}$ is Zariski dense in $\mathfrak{u}$ in case $\mathfrak{u}\cap\mathfrak{r}$ is not empty. Theorem 14 also generalizes most of Theorem 10 (case where U has only one point).

THEOREM 14. *Let* $J^U\subseteq J$, $U\in\mathcal{U}$, *be any prime ideal in* J. *Then* J^US *is a prime ideal in* S *and*

$$(4.10.1)\qquad \mathfrak{u}(J^US)=\bigcup_{\xi\in U}P(\xi).$$

That is, $\mathfrak{u}(J^U S) = u^{-1}(U)$ *where* u *is the map* (3.3.2). *Moreover*

$$U \to \mathfrak{u}(J^U S) \tag{4.10.2}$$

defines a one-one correspondence between the set of all Zariski closed subvarieties of $\boldsymbol{C}^l$ *and the set of all* G*-stable Zariski closed subvarieties* $\mathfrak{u} \subseteq \mathfrak{g}$ *such that* $\mathfrak{u} \cap \mathfrak{r}$ *is not empty. In particular* $U = u(\mathfrak{u})$ *is in* $\mathcal{U}$ *for such a subvariety* $\mathfrak{u} \subseteq \mathfrak{g}$, $\mathfrak{u} = \mathfrak{u}(J^U S)$ *and*

$$\mathfrak{u} \cap \mathfrak{r} = \bigcup_{\xi \in U} O^{\mathfrak{r}}(\xi). \tag{4.10.3}$$

Let $U \in \mathcal{U}$ *and put* $\mathfrak{u} = \mathfrak{u}(J^U S)$. *Then*

$$\operatorname{codim} U \text{ in } \boldsymbol{C}^l = \operatorname{codim} \mathfrak{u} \text{ in } \mathfrak{g}. \tag{4.10.4}$$

Furthermore if $x \in \mathfrak{u} \cap \mathfrak{r}$ *then* x *is a simple point of* $\mathfrak{u}$ *if and only if* $u(x)$ *is a simple point of* U. *Finally if* $\mathfrak{r}^o$ *is the* (*Zariski closed in* $\mathfrak{g}$) *complement of* $\mathfrak{r}$ *in* $\mathfrak{g}$ *then*

$$\operatorname{codim} \mathfrak{u} \cap \mathfrak{r}^o \text{ (in } \mathfrak{u}) \geqq 2. \tag{4.10.5}$$

Proof. Let J' be any ideal in J. It is immediate that $J'S$ is the image of $J' \otimes H$ under the isomorphism (4.8.4) and hence one has

$$J'S \cap J = J'. \tag{4.10.6}$$

Now assume that J' is a radical ideal (an ideal equal to its own radical) in J. We will show that $J'S$ is a radical ideal in S.

Let J'_* be the radical ideal in $S(\boldsymbol{C}^l)$ corresponding to J' under the isomorphism $J \to S(\boldsymbol{C}^l)$ where $f \to f^*$ and let $U \subseteq \boldsymbol{C}^l$ be the Zariski closed set of all $\xi \in \boldsymbol{C}^l$ at which J'_* vanishes. It is obvious that if $\mathfrak{u}$ is the Zariski closed set, in $\mathfrak{g}$, of all $x \in \mathfrak{g}$ at which $J'S$ vanishes then $\mathfrak{u} = u^{-1}(U)$ or

$$\mathfrak{u} = \bigcup_{\xi \in U} P(\xi). \tag{4.10.7}$$

To prove $J'S$ is a radical ideal it suffices to show that if $f \in S$ is assumed to vanish on $\mathfrak{u}$ then $f \in J'S$. By Theorem 11, § 4.8, we can write $f = \sum f_i h_i$ where $f_i \in J$, $h_i \in H$ and the h_i are linearly independent. Let $\xi \in U$. Then since the f_i reduce to constants on $P(\xi)$ it follows from the isomorphism (4.8.6) and (3.8.7) that since f vanishes on $P(\xi)$ the f_i also vanish on $P(\xi)$. Thus the f_i are in J' by the nullstellensatz and hence $f \in J'S$ so that $J'S$ is a radical ideal.

Now let $U \in \mathcal{U}$ so that J^U is prime in J. Put $J' = J^U$ so that, from above, $J^U S$ is a radical ideal in S. To prove that $J^U S$ is prime it suffices now only to show that $\mathfrak{u}$ is irreducible.

Let $f \in S$ and let $U(f)$ be the set of all $\xi \in U$ such that $f \mid P(\xi)$ is not zero. Obviously $f \mid \mathfrak{u}$ is not zero if and only if $U(f)$ is not empty. We first show that in such a case $U(f)$ contains a non-empty Zariski open subset of U. Indeed assume $U(f)$ is not empty and $\xi \in U(f)$. Then $f \mid O^{\mathfrak{r}}(\xi)$ is not zero by (3.8.7). Hence there exists $a \in G$ such that $(a \cdot f) \mid O^{\mathfrak{r}}(\xi) \cap \mathfrak{v}$ is not zero, by Theorem 8, § 4.7, where $\mathfrak{v}$ is defined as in (4.5.6). Thus $(a \cdot f) \mid \mathfrak{v}$ does not vanish on a Zariski subset of $\mathfrak{v}$ containing $O^{r}(\xi) \cap \mathfrak{v}$. Using the isomorphism (4.7.3) it follows that $U(a \cdot f)$ contains a non-empty Zariski open subset of U. But clearly $U(a \cdot f) = U(f)$. Hence $U(f)$ contains such a subset.

Now let $f_i \in S$, $i = 1, 2$, be arbitrary except that $f_i \mid \mathfrak{u}$ is not zero. To show $\mathfrak{u}$ is irreducible we must show that $f_1 f_2 \mid \mathfrak{u}$ is not zero. From above it follows that $U(f_i)$ contain a non-empty Zariski open subset of U. But since U is irreducible it follows that $U(f_1) \cap U(f_2)$ is not empty. But then $f_1 f_2 \mid P(\xi)$ is not zero in case $\xi \in U(f_1) \cap U(f_2)$ since $P(\xi)$ is irreducible by Theorem 3, § 3.8. Thus $\mathfrak{u}$ is irreducible and hence $J^U S$ is prime.

The relation (4.10.1) is just (4.10.7). Furthermore if $\mathfrak{u} = \mathfrak{u}(J^U S)$ then (4.10.3) follows from (3.8.4).

Moreover, using (4.10.1), it is immediate that the map, given by (4.10.2), from $\mathfrak{U}$ into the set of all Zariski closed G-stable subvarieties $\mathfrak{u}$ of $\mathfrak{g}$ such that $\mathfrak{u} \cap \mathfrak{r}$ is not empty is injective. Now assume that $\mathfrak{u}$ is such a subvariety. We will show that $\mathfrak{u}$ is in the image of the map defined by (4.10.2). Let the set $U \subseteq \boldsymbol{C}^l$ be defined by putting $U = u(\mathfrak{u} \cap \mathfrak{r})$. Since $\mathfrak{u}$ is Zariski irreducible and $\mathfrak{u} \cap \mathfrak{r}$ is Zariski dense in $\mathfrak{u}$ it follows that U is Zariski irreducible. On the other hand by Theorem 8, § 4.7, it is clear that U corresponds to $\mathfrak{u} \cap \mathfrak{v}$ under the isomorphism (4.7.3). But since $\mathfrak{u} \cap \mathfrak{v}$ is Zariski closed in $\mathfrak{v}$ it follows that U is Zariski closed in $\boldsymbol{C}^l$. Hence $U \in \mathfrak{U}$. But U is Zariski dense in $u(\mathfrak{u})$. But this implies $u(\mathfrak{u}) = U$ since U is Zariski closed. Thus $\mathfrak{u} \subseteq u^{-1}(U)$. But $u^{-1}(U)$ is clearly in the Zariski closure of $\mathfrak{u} \cap \mathfrak{r}$ since the relation (4.10.3) obviously holds. Thus $\mathfrak{u} = u^{-1}(U)$ or $\mathfrak{u} = \mathfrak{u}(J^U S)$.

Now obviously $(du_i^*)_\xi$, $i = 1, 2, \cdots, l$, are linearly independent at any point $\xi \in \boldsymbol{C}^l$. Since $(df)_x$ is in the span of the $(du_i)_x$ for any $f \in J$ and $x \in \mathfrak{g}$ it follows from Theorem 9, § 4.7, that $(df)_x = 0$ if and only if $(df^*)_\xi = 0$ for any $f \in J$ and $x \in \mathfrak{r}$, where $\xi = u(x)$. It follows in particular that if $U \in \mathfrak{U}$ and $x \in \mathfrak{u} \cap \mathfrak{r}$ where $\mathfrak{u} = \mathfrak{u}(J^U S)$ then the dimension r_x of the space spanned the $(df)_x$ for all $f \in J^U$ is the same as the dimension r_ξ of the space spanned by all $(df^*)_\xi$ where $f \in J^U$ and $\xi = u(x)$. If r is the codimension of

U in C^l and U_s is the set of simple points of U then by the Zariski criterion $r_\xi \leqq r$ for all $\xi \in U$ and $r_\xi = r$ if and only if $\xi \in U_s$. Thus $r_x \leqq r$ for all $x \in \mathfrak{u} \cap \mathfrak{r}$ and $r_x = r$ if and only if $u(x) \in U_s$. Since $\mathfrak{u} \cap \mathfrak{r}$ is Zariski open in $\mathfrak{u}$ and since J^U generates the prime ideal of $\mathfrak{u}$ then by applying the Zariski criterion to $\mathfrak{u}$ one obtains (4.10.4) and the fact that $x \in \mathfrak{u} \cap \mathfrak{r}$ is simple on $\mathfrak{u}$ if and only if $u(x)$ is simple on U.

Finally let $U \in \mathcal{U}$ have codimension r in C^l. Let $\mathfrak{a}$ be an irreducible component of $\mathfrak{u} \cap \mathfrak{r}^c$ where $\mathfrak{u} = \mathfrak{u}(J^U S)$ and let $Y \subseteq U$ be the Zariski closure of $u(\mathfrak{a})$. Let $q = \dim Y$ so that, obviously, $q \leqq l - r$. But if $p = \dim \mathfrak{a}$ then by Corollary 1, p. 109 in [3] there exists $\xi \in Y$ such that $(u \mid \mathfrak{a})^{-1}(\xi) = P(\xi) \cap \mathfrak{a}$ has dimension $p - q$. Hence $\dim P(\xi) \cap \mathfrak{a} \geqq p - l + r$. But by (3.8.5) one has $\dim P(\xi) \cap \mathfrak{r}^c \leqq n - l - 2$. Thus $n - 2 \geqq p + r$ or $n - r \geqq p + 2$. That is, by (4.10.4) $\dim \mathfrak{u} \geqq \dim \mathfrak{a} + 2$. This proves (4.10.5). Q.E.D.

Remark 23. By putting $\mathfrak{u} = \mathfrak{g}$ in (4.10.5) note that $\mathfrak{r}^c$ has a dimension of at least 2 in $\mathfrak{g}$. On the other hand if $\mathfrak{q}$ is the set of regular elements in $\mathfrak{g}$ and $v \in J$ is the invariant polynomial such that $v \mid \mathfrak{h}$ is the product of all the roots (positive as well as negative, so that it is a Weyl group invariant) then the complement $\mathfrak{q}^c$ of $\mathfrak{g}$ is the set of zeros of v and hence has codimension 1 in $\mathfrak{g}$.

5. The normality of the varieties $P(\xi)$ and the generalized exponents. The following criterion for normality and its proof is due to Seidenberg.

THEOREM 15 (Seidenberg). *Let $\mathfrak{u} \subseteq \mathfrak{g}$ be a Zariski closed subvariety of $\mathfrak{g}$. Assume* (a) *that $\mathfrak{u}$ is a complete intersection and* (b) *the set of non-simple point of $\mathfrak{u}$ has a codimension of at least two in $\mathfrak{u}$. Then $\mathfrak{u}$ is a normal variety.*

Proof. Let $r = \dim \mathfrak{u}$. By [16], Theorem 3, one knows that $\mathfrak{u}$ is normal if (1) $\mathfrak{u}$ is free of $(r-1)$-dimensional singularities and (2), every principal ideal in the affine algebra of $\mathfrak{u}$ is unmixed. Since assumption (1) is satisfied (statement (b) in Theorem) it suffices therefore only to show that if $I(\mathfrak{u})$ is the prime ideal, in S, corresponding to $\mathfrak{u}$ and $f \in S$ then the ideal $(I(\mathfrak{u}), f)$ is unmixed. Obviously one may assume that $(I(\mathfrak{u}), f)$ has dimension $r - 1$. But then the result follows from Macaulay's theorem (see [19], p. 203) since, by (a), one has that $I(\mathfrak{u}) = (f_1, \cdots, f_{n-r})$ for some $f_i \in I(\mathfrak{u})$. Q.E.D.

Now if V is any finite dimensional G-module and $x \in \mathfrak{g}$ is arbitrary consider V^{G^x} the subspace of vectors in V that are fixed under all elements of G^x. We now find that the dimension of V^{G^x} is the multiplicity of the zero weight in V (and hence is the same) for all $x \in \mathfrak{r}$. (This, incidentally is not

necessarily true for the covering group of G). Obviously it is enough to show this for irreducible G-modules.

THEOREM 16. *Let $\xi \in \boldsymbol{C}^l$ be arbitrary. Then $P(\xi)$ is a normal variety. That is, if $x \in \mathfrak{r}$ then $\bar{O}_x$ is a normal variety (see Theorem 4, § 3.8). Furthermore if $R(O_x)$ denotes the ring of everywhere defined functions on the orbit O_x ($R(O_x)$ is isomorphic to $R(G/G^x)$) then $R(O_x)$ is an affine algebra. (That is, it is finitely generated.) In fact $R(O_x) = S(O_x)$ where $S(O_x)$ is the restriction of S to O_x so that if $\xi = u(x)$ then $P(\xi) = \bar{O}_x$ is the affine variety of maximal ideals of $R(O_x)$.*

Moreover $R(O_x)$ is a completely reducible G-module and for any $\lambda \in D$ the multiplicity of ν^λ in $R(O_x)$ is l_λ, where l_λ is the multiplicity of the zero weight of ν_λ (or equally of ν^λ), so that $R(O_x)$, for all $x \in \mathfrak{r}$, are isomorphic as G-modules. Finally for any $x \in \mathfrak{r}$

$$\dim V_\lambda{}^{G^x} = l_\lambda \tag{5.1.1}$$

where V_λ and ν_λ is defined as in § 2.1.

Proof. By Theorem 10, § 4.8, $P(\xi)$ is a complete intersection. Also by Theorem 10 the set of non-simple points of $P(\xi)$ has a codimension of at least two in $P(\xi)$. Hence $P(\xi)$ is normal by Theorem 15. But now by Corollary 1, § 3.8, the complement of O_x in $\bar{O}_x$ has a codimension of at least two in $\bar{O}_x$ for any $x \in r$. Hence $R(O_x) = S(O_x)$ by Proposition 9, § 2.2. But in any case $S(O_x)$ is isomorphic to $S(\bar{O}_x)$. Since $S(\bar{O}_x) = R(\bar{O}_x)$ (see (2.2.2)) it follows that every element of $R(O_x)$ extends uniquely to an element of $R(\bar{O}_x)$. This induces an isomorphism

$$R(O_x) \to R(\bar{O}_x) \tag{5.1.2}$$

so that $R(O_x)$ is an affine algebra. Obviously $R(O_x)$ is a completely reducible G-module. But by Theorem 11, § 4.8, the multiplicity of ν^λ in $S(O_x)$ is l_λ. It follows therefore from Proposition 8, § 2.1, and the G-module isomorphism (2.2.1) that $\dim V_\lambda{}^{G^x} = l_\lambda$. Q. E. D.

Let $x \in \mathfrak{r}$. As a corollary of Theorem 16 we now observe that $\bar{O}_x$ is distinguished among all affine varieties into which O_x may be embedded.

COROLLARY 3. *Let $x \in \mathfrak{r}$ be arbitrary so that $\xi = u(x) \in \boldsymbol{C}^l$ is arbitrary. If we identify G/G^x with O_x (using the isomorphism (1.2.2)) then any morphism of G/G^x into any affine variety X extends uniquely to a morphism of the affine variety $P(\xi)$ into X.*

In particular any morphism of the orbit of principal nilpotent elements

into an affine variety X extends to a morphism of the variety of all nilpotent elements into X.

Proof. Using the isomorphism (5.1.2) this result follows from Corollary 1, p. 58 in [3]. Q.E.D.

Remark 24. By Theorem 16 all $R(O_x)$ for $x \in \mathfrak{r}$, are isomorphic as G-modules. However it should be noted that they are not isomorphic as rings. Indeed the corresponding variety of maximal ideals of $R(O_x)$ is non-singular in case x is regular and, by Theorem 10, § 4.8, is singular otherwise (for example, in case x is principal nilpotent).

5.2. Now for any $\lambda \in D$ consider $\mathrm{Hom}_G(V^\lambda, S)$ the space of all G-module maps γ of V^λ into the ring of polynomials S. Obviously $\gamma(V^\lambda) \subseteq S^\lambda$ for any such γ.

We now observe that any $x \in \mathfrak{g}$ induces a linear map

$$\omega_x\colon \mathrm{Hom}_G(V^\lambda, S) \to V_\lambda{}^{G^x}$$

by the relation

$$\langle v, \omega_x\gamma\rangle = (\gamma(v))(x) \tag{5.2.1}$$

for all $v \in V^\lambda$ and any $\gamma \in \mathrm{Hom}_G(V^\lambda, S)$. Since γ is a G-module map it follows immediately from (1.1.5) that for any $a \in G$

$$\nu_\lambda(a)\omega_x(\gamma) = \omega_{ax}(\gamma) \tag{5.2.2}$$

and hence, obviously, $\omega_x(\gamma) \in V_\lambda{}^{G^x}$ for any $x \in \mathfrak{g}$.

Now by Theorem 11, § 4.8, the subspace $\mathrm{Hom}_G(V^\lambda, H)$ of $\mathrm{Hom}_G(V^\lambda, S)$ is of dimension l_λ. On the other hand, by (5.1.1), $V_\lambda{}^{G^x}$ is also l_λ-dimensional whenever $x \in \mathfrak{r}$. As a corollary of Theorem 16 we obtain

COROLLARY 4. *Let $x \in \mathfrak{r}$ and let $\lambda \in D$. Then the map*

$$\mathrm{Hom}_G(V^\lambda, H) \to V_\lambda{}^{G^x} \tag{5.2.3}$$

defined by restricting ω_x is an isomorphism.

Proof. Since both sides of (5.2.3) are vector spaces of dimension l_λ it suffices to show (5.2.3) is a monomorphism. Let γ be in the kernel of (5.2.3). Then, by (5.2.2), $\omega_{ax}(\gamma) = 0$ for all $a \in G$. Now let $v \in V^\lambda$. Then, by (5.2.1), $(\gamma(v))(ax) = 0$ for all $a \in G$. Thus $\gamma(v) \mid O_x = 0$. But since (4.8.6) is an isomorphism it follows that $\gamma(v) = 0$ for all $v \in V^\lambda$. Hence $\gamma = 0$, and consequently (5.2.3) is an isomorphism. Q.E.D.

5.3. Now since ad maps $\mathfrak{g}$ into the Lie algebra of G it is clear that any finite dimensional representation

$$\nu\colon G \to \mathrm{Aut}\, V$$

of G induces, by taking differentials, a representation of $\mathfrak{g}$, which we also denote by ν. Note that one always has $\nu(\mathfrak{z})=0$ where $\mathfrak{z}$ is the center of $\mathfrak{g}$.

Now let Z be the subgroup of $\mathfrak{h}'$ generated by the set of roots Δ. If $o(\mu)$, the order of μ, is defined by

$$o(\mu)=\langle x_0,\mu\rangle$$

for any $\mu\in Z$ it is then clear from (4.1.6) that $o(\mu)$ is always an integer. Now since every weight of ν is clearly necessarily an element of Z it follows therefore that

$$V=\sum_{k\in Z} V^{(k)}$$

is a direct sum where $V^{(k)}$ is the eigenspace of $\nu(x_0)$ belonging to the eigenvalue k. Obviously

$$\nu(\mathfrak{g}^{(i)})V^{(k)}\subseteq V^{(i+k)}. \tag{5.3.1}$$

Now e_- be the principal nilpotent element defined as in § 4.2. For notational simplicity write $F=V^{G^{e_-}}$. Since G^{e_-} is connected (Proposition 14, § 3.6) one also has

$$F=\operatorname{Ker}\nu(\mathfrak{g}^{e_-}) \tag{5.3.2}$$

and by (5.1.1)

$$\dim F=l_\nu$$

where l_ν is the dimension of the zero weight of ν.

Now since x_0 lies in the normalizer of $\mathfrak{g}^{e_-}$ it follows from (5.3.2) that F is stable under $\nu(x_0)$. But then we observe that there exists a unique sequence of integers $m_i(\nu)$, $i=1,2,\cdots,l_\nu$, where

$$m_1(\nu)\leqq\cdots\leqq m_{l_\nu}(\nu)$$

such that F has a basis v_i, $i=1,2,\cdots,l_\nu$, where

$$v_i\in V^{(-m_i(\nu))}. \tag{5.3.3}$$

Remark 25. By applying the inner automorphism which carries e_+ into e_- and x_0 into $-x_0$ note that we would get the same integers $m_i(\nu)$ by using e_+ into instead of e_- and dropping the minus in (5.3.3).

Observe then that the $m_i(\nu)$ generalizes the notion of exponents. Indeed if ν is the adjoint representation then $l_\nu=l$ and, by Theorem 5, § 4.4, $m_i(\nu)=m_i$ since $F=\mathfrak{g}^{e_-}$.

Since $F\subseteq\operatorname{Ker}\nu(e_-)$ it follows from the representation theory of a three dimensional simple Lie algebra (e.g. see [13], § 2.5) that for any i

$$0\leqq m_i(\lambda). \tag{5.3.4}$$

Now, as in Remark 22, § 4.8, identify D with the (subset of Z) set of all dominant integral (with respect to G) forms on $\mathfrak{h}$ so that any $\lambda \in D$ is the highest weight of ν^λ. Note then that $-\lambda$ is the lowest weight of ν_λ.

When $V = V_\lambda$ and $\nu = \nu_\lambda$ we will write F_λ for F and $m_i(\lambda)$ for $m_i(\nu_\lambda)$. The $m_i(\lambda)$, $i = 1, 2, \cdots, l_\lambda$, will be called the generalized exponents of $\mathfrak{g}$ (corresponding to λ). See Remark 25.

Now let $v_\lambda \in V_\lambda$ be the lowest weight $(-\lambda)$ vector. Then since V_λ, as one knows, is a cyclic module with respect to the universal enveloping algebra of $\mathfrak{m}$ with v_λ as cyclic vector it follows that

$$(V_\lambda)^{(-k)} = \begin{cases} 0 \text{ if } k > o(\lambda) \\ (v_\lambda) \text{ if } k = o(\lambda). \end{cases}$$

Since $V_\lambda^{(-o(\lambda))}$ is obviously contained in F_λ. (One uses the relation $\mathfrak{g}^{e-} \subseteq \mathfrak{m}^* + \mathfrak{z}$ mentioned in the proof of Theorem 5, § 4.4). It follows then that

$$m_i(\lambda) < m_{l_\lambda}(\lambda) = o(\lambda) \tag{5.3.5}$$

for $1 \leqq i < l_\lambda$.

5.4. Let $\lambda \in D$. For convenience we will write $H(\lambda)$ and $S(\lambda)$ for the subspaces $H^\lambda \subseteq H$ and $S^\lambda \subseteq S$ respectively. See § 2.1. Clearly $H(\lambda)$ and $S(\lambda)$ are graded subspaces. In particular then

$$H(\lambda) = \sum_{j=0}^{\infty} H(\lambda)^j.$$

On the other hand by Theorem 11, § 4.8, the multiplicity of ν^λ in H is l_λ. Hence one has a direct sum

$$H(\lambda) = \sum_{i=1}^{l_\lambda} H_i(\lambda) \tag{5.4.1}$$

where (1) $H_i(\lambda)$ is an irreducible G-module, (2) $H_i(\lambda)$ is a space of homogeneous polynomials so that for some degree $n_i(\lambda)$ one has $H_i(\lambda) \subseteq S^{n_i(\lambda)}$ where (3) we may assume the $n_i(\lambda)$ are monotone non-decreasing with i.

The integers $n_i(\lambda)$ are the degrees k such that ν^λ occurs in H^k. The question arises: how does one determine these integers? Since $S(\lambda)$ is obviously isomorphic to the tensor product $J \otimes H(\lambda)$ by Theorem 11, § 4.8, it is clear that such information is needed if one is to determine the formal power series

$$q_\lambda(t) = \sum_{k=0}^{\infty} \dim S(\lambda)^k t^k$$

and, as a consequence, the multiplicity of ν^λ in S^k for any k.

The following theorem asserts that the $n_i(\lambda)$ are exactly the generalized exponents $m_i(\lambda)$.

THEOREM 17. *For any $\lambda \in D$ and $i = 1, 2, \cdots, l_\lambda$, one has $n_i(\lambda) = m_i(\lambda)$ so that $H_i(\lambda) \subseteq S^{m_i(\lambda)}$. In particular $k = o(\lambda)$ is the maximum degree such that $H(\lambda)^k \neq 0$. Furthermore $H(\lambda)^k$ is irreducible for this value of k. That is,*

$$H_{l_\lambda}(\lambda) = H(\lambda)^{o(\lambda)}.$$

Moreover the formal power series $q_\lambda(t)$ may be given by

$$q_\lambda(t) = d_\lambda \frac{\sum_{i=1}^{l_\lambda} t^{m_i(\lambda)}}{\prod_{i=1}^{l} (1 - t^{m_i})} \tag{5.4.2}$$

where $d_\lambda = \dim V_\lambda$.

Proof. Let v'_i, $i = 1, 2, \cdots, l_\lambda$, be a basis of F_λ such that $v'_i \in V_\lambda^{(-m_i(\lambda))}$. Now let $c \in \boldsymbol{C}^*$ be arbitrary. Let $r \in \boldsymbol{C}$ be such that $e^{-r} = c$ and let $a \in G$ be defined by putting $a = \exp r \operatorname{ad} x_0$. It is then clear that

$$\nu_\lambda(a) v'_i = c^{m_i(\lambda)} v'_i. \tag{5.4.3}$$

Also note that

$$a(e_-) = c e_-. \tag{5.4.4}$$

Now by Corollary 14, § 5.2, there exists a basis γ_i, $i = 1, 2, \cdots, l_\lambda$, of $\operatorname{Hom}_G(V^\lambda, H)$ such that

$$\omega_{e_-}(\gamma_i) = v'_i.$$

But now by (5.4.4-5) and (5.2.2) one gets the equation

$$c^{m_i(\lambda)} \omega_{e_-}(\gamma_i) = \omega_{ce_-}(\gamma_i). \tag{5.4.5}$$

Substituting in (5.2.1) this implies that for any $v \in V^\lambda$

$$(\gamma_i(v))(ce_-) = c^{m_i(\lambda)} \gamma_i(v)(e_-).$$

But then, conjugating by G (and using (1.1.5)) it also follows that

$$\gamma_i(v)(cy) = c^{m_i(\lambda)} \gamma_i(v)(y)$$

for all $y \in O_{e_-}$. But then since (4.8.6) is an isomorphism for $x = e_-$ it follows easily by choosing c, for example, to be positive that

$$\gamma_i(V^\lambda) \subseteq S^{m_i(\lambda)}.$$

On the other hand by definition of the γ_i

$$H(\lambda) = \sum_{i=1}^{l_\lambda} \gamma_i(V^\lambda)$$

is a direct sum of irreducible G-modules. By uniqueness of the $n_i(\lambda)$ it then follows that $n_i(\lambda) = m_i(\lambda)$ for $i = 1, \cdots, l_\lambda$.

The second and third statements of the theorem follow immediately from (5.3.5). The equation (5.4.2) is an immediate consequence of the obvious fact that the isomorphism (4.8.4) induces an isomorphism of $J \otimes H(\lambda)$ onto $S(\lambda)$. Q.E.D.

Remark 26. We observe here that Theorem 17 is a generalization of Theorem 5, § 4.4, asserting that $m_i(\nu) = m_i$ where ν is the adjoint representation. Indeed let U be the subspace of all $u \in J^+$ such that $\langle \partial, u \rangle = 0$ where $\partial \in (J_*^+)^2$. It follows immediately that $\dim U = l$ and that any homogeneous basis of U is a set of primitive invariants u_i, $i = 1, 2, \cdots, l$. But if the u_i are so chosen then by definition of U one has $\partial u_i \in J^0 = S^0$ (immediate from (1.1.2)) for any $\partial \in J_*^+$. It follows immediately then that, for $1 \leqq i \leqq l$,

$$(5.4.6) \qquad \partial_x u_i \in H^{m_i} \text{ for every } x \in \mathfrak{g}.$$

(Recall that S_* is commutative.) But now if g_1 is a simple component of $\mathfrak{g}$ of rank l_1 and u_{i_k}, $k = 1, 2, \cdots, l_1$, are as in the proof of Theorem 9, § 4.7, then one must have $\partial_x u_{i_k} \in H(\psi)$ for any $x \in \mathfrak{g}_1$ where $\psi \in D$ is defined as in the proof of Theorem 9. See (4.7.8). (In particular note that if $\mathfrak{g}$ is simple and x_j is a basis of $\mathfrak{g}$ then

$$(5.4.7) \qquad \partial_{x_j} u_i,\ i = 1, \cdots, l,\ j = 1, \cdots, n \text{ is a basis of } H(\psi)$$

where ψ is the highest root of $\mathfrak{g}$). It follows immediately that $n_k(\psi) = m_{i_k}$. Applying Theorem 17 one then obtains $m_k(\psi) = m_{i_k}$ which immediately yields $m_j(\nu) = m_j$.

It should also be observed then that the relation $o(\lambda) = m_{l_\lambda}(\lambda)$ generalizes the well known relation $o(\psi) = m_l$. See Corollary 8.6 and Lemma 9.1 in [13].

Remark 27. Note that the argument given in the proof of Theorem 17 may be reversed. That is, if L is an arbitrary irreducible G-submodule of $H(\lambda)$ and $\gamma \in \operatorname{Hom}_G(V^\lambda, H)$ is such that

$$(5.4.8) \qquad \gamma(V^\lambda) = L$$

then for any integer $j \geqq 0$ one has

$$(5.4.9) \qquad L \subseteq S^j \text{ if and only if } \omega_{e_-}(\gamma) \in V_\lambda^{(-j)}.$$

Indeed if $L \subseteq S^j$, then by (5.2.1), $\omega_{ce_-}(\gamma) = c^j \omega_{e_-}(\gamma)$ for any $c \in \boldsymbol{C}^*$. But then if $a \in G$ is defined as in the proof of Theorem 17 one obviously has

$\omega_{e_-}(\gamma) \in V_\lambda^{(-j)}$ since $\nu_\lambda(a)\omega_{e_-}(\gamma) = c^j\omega_{e_-}(\gamma)$, by (5.4.4) and (5.2.2) and $c \in \boldsymbol{C}^*$ is arbitrary. The argument for the other direction has been given in the proof of Theorem 17.

For any $\lambda \in D$ let $\gamma_i^\lambda \in \operatorname{Hom}_G(V^\lambda, H)$, $i = 1, 2, \cdots, l_\lambda$ be fixed so that $\gamma_i^\lambda(V^\lambda) = H_i(\lambda)$. Obviously the γ_i^λ are a basis of $\operatorname{Hom}_G(V^\lambda, H)$ by (5.4.1).

Thet argument in Remark 27 may be used to yield

THEOREM 18. *Let $x \in \mathfrak{r}$ and $a \in G$. Assume $a(x) = cx$ for some $c \in \boldsymbol{C}^*$. Now let $\lambda \in D$. Then the l_λ-dimensional* (by 5.1.1)) *space $V_\lambda^{G^e}$ is stable under $\nu_\lambda(a)$. Furthermore for $i = 1, 2, \cdots, l_\lambda$,*

$$\{c^{m_i(\lambda)}\} \text{ are the eigenvalues of } \nu_\lambda(a) \mid V_\lambda^{G^e}$$

and $\omega_e(\gamma_i^\lambda)$ is a corresponding basis of eigenvectors.

Proof. Since $H_i(\lambda) \subseteq S^{m_i(\lambda)}$ by Theorem 17 it follows from (5.2.1) that $\omega_{ce}(\gamma_i^\lambda) = c^{m_i(\lambda)}\omega_e(\gamma_i^\lambda)$. But then $\nu_\lambda(a)\omega_e(\gamma_i^\lambda) = c^{m_i(\lambda)}\omega_e(\gamma_i^\lambda)$ by (5.2.2). On the other hand by Corollary 4, § 5.2, the $\omega_e(\gamma_i^\lambda)$ are a basis of $V_\lambda^{G^e}$.

Q. E. D.

5.5. Let V be a finite dimensional G-module with respect to a representation ν. Let $A \subseteq G$ be the Cartan subgroup of G corresponding to $\mathfrak{h}$ so that V^A is the zero weight space. We recall that W is the Weyl group of G corresponding to $\mathfrak{h}$. Now since V^A is obviously stable under the normalizer of A in G it follows that ν induces a representation

$$\pi: W \to \operatorname{Aut} V^A$$

of the Weyl group W on V^A. One notes that this is a generalization of the usual representation of W on $\mathfrak{h}$ (case where ν is the adjoint representation).

When $V = V_\lambda$ and $\nu = \nu_\lambda$ we will write π_λ for π.

Now assume in the remainder of this section that $\mathfrak{g}$ is a non-trivial simple Lie algebra. Let $\psi \in D$ be the highest root of $\mathfrak{g}$ so that ν^ψ is the adjoint representation. Let

$$s = 1 + o(\psi)$$

or equivalently let $s = 1 + m_l$. See end of Remark 26, § 5.4.

We recall that an element $\sigma \in W$ is called a Coxeter-Killing transformation in [13], § 8.1, if it can be expressed as the product of the reflections defined by the simple roots (in any order) relative to any system of positive roots.

Let $\sigma \in W$ be a Coxeter-Killing transformation. It was observed empirically by Coxeter and then proved independently by Steinberg and in [13]

that the order of σ is s. It was also observed empirically by Coxeter and proved in [4] that the eigenvalues of $\pi_\psi(\sigma)$ are $e^{2\pi i m_j/s}$, $j=1,2,\cdots,l$. We will now generalize this for all $\lambda\in D$.

By a theorem of Coleman (see [4]) there exists a regular element $z\in\mathfrak{h}$ such that

$$\sigma(z)=e^{2\pi i/s}z. \tag{5.5.1}$$

In fact, by Corollary 9.2 in [13], z is a cyclic element of $\mathfrak{g}$.

THEOREM 19. *Let $\sigma\in W$ be a Coxeter-Killing transformation and let $\lambda\in D$ be arbitrary. Then the eigenvalues of $\pi_\lambda(\sigma)$ are $e^{2\pi i m_i(\lambda)/s}$, $i=1,2,\cdots,l_\lambda$, and if $z\in\mathfrak{h}$ is the cyclic element satisfying* (5.5.1) *then $\{\omega_z(\gamma_i^\lambda)\}$ is a corresponding basis of eigenvectors.*

Proof. Let $a\in G$ be any element of the normalizer of A which defines $\sigma\in W$ so that $\nu_\lambda(a)\mid V_\lambda^A=\pi_\lambda(\sigma)$. Since $a(z)=cz$ where $c=e^{2\pi i/s}$ the result follows immediately from Theorem 18 since by Coleman's theorem z is regular and hence $z\in\mathfrak{r}$. Q.E.D.

Remark 28. If $\mathfrak{g}$ is the three dimensional simple Lie algebra we may identify D with the set of non-negative integers where $\dim V^\lambda=2\lambda+1$. Here $l_\lambda=1$ for all $\lambda\in D$ and $m_1(\lambda)=\lambda$ by (5.3.5) since $o(\lambda)=\lambda$. Also obviously $s=2$. Note then that one recovers from Theorem 19 the well known fact that $\pi_\lambda(\sigma)=(-1)^\lambda$ for $\sigma\in W$, $\sigma\neq 1$.

Remark 29. For any $k\in\boldsymbol{Z}$ let $[k]$ denote its canonical image in $\boldsymbol{Z}_s=\boldsymbol{Z}/s\boldsymbol{Z}$ and for any $m\in\boldsymbol{Z}$ let $r_\lambda(m)$ be the number of integers $1\leq i\leq l_\lambda$ such that $[m_i(\lambda)]=[m]$. Now if $\sigma\in W$ is a Coxeter-Killing transformation and $k\in\boldsymbol{Z}$ is prime to s then σ^k is again a Coxeter-Killing transformation by Corollary 9.2 in [13] (using (5.5.1)). It follows easily therefore from Theorem 19 that $r_\lambda(m)=r_\lambda(km)$ for any integer m. Note that this generalizes Chevalley's observation that $m_i+m_{l+1-i}=s$ since we may put $k=-1$ and one knows $m_i<s$.

6. A decomposition theorem for the universal enveloping algebra U of $\mathfrak{g}$. 1. Let T be the tensor algebra over $\mathfrak{g}$. Then one knows that in a unique way T is a G-module so that G operates as a group of algebra automorphisms extending the action of G on $\mathfrak{g}$.

Let $Q\subseteq T$ denote the G-submodule of all symmetric tensors in T.

Also for $\tau=0,1$ let I_τ be the ideal in T generated by all

$$(x\otimes y-y\otimes x)-\tau[x,y]$$

where $x,y\in\mathfrak{g}$. Then since I_τ is G-stable the algebra $T_\tau=T/I_\tau$ is a G-module

and one knows that the canonical epimorphism $T \to T_\tau$ induces a G-module isomorphism

$$\delta_\tau : Q \to T_\tau.$$

However, by definition the G-module T_0 is the symmetric algebra S_* with the G-module structure of § 1.1 and the G-module T_1 is the universal enveloping algebra U of $\mathfrak{g}$ with G operating by the usual extension of the adjoint representation. Now if $\delta_* = \delta_1 \circ \delta_0^{-1}$ then obviously

$$\delta_* : S_* \to U$$

is a G-module isomorphism and that furthermore since Q is the subspace of T generated by all tensors of the form $x \otimes \cdots \otimes x$ where $x \in \mathfrak{g}$ it follows that $\delta_*((\partial_x)^k) = x^k$ for any $x \in \mathfrak{g}$. Finally then if we compose δ_* with the inverse of B (see § 1.4) one obtains a G-module isomorphism

$$\delta : S \to U$$

such that, for any integer $k \geqq 0$,

$$\delta(g^k) = (\delta(g))^k \tag{6.1.1}$$

for every $g \in S^1$. Furthermore, one knows, by the theorem of Birkhoff-Witt that the filtration of U defined by the G-submodules

$$U_k = \delta(\sum_{i=0}^{k} S^i)$$

is such that

$$\delta(fg) = \delta(f)\delta(g) \bmod U_{i+j-1} \tag{6.1.2}$$

for (not necessarily homogeneous) polynomials f and g where $\deg f \leqq i$ and $\deg g \leqq j$. In particular (6.1.2) implies

$$U_i U_j \subseteq U_{i+j}. \tag{6.1.3}$$

Now let $Z \subseteq U$ be the center of U. By a theorem of Chevalley one knows that Z, like J, is isomorphic to a polynomial ring in l generators. Furthermore since Z is clearly the subalgebra of fixed elements under the action of G on U and since δ is a G-module isomorphism it follows that

$$\delta(J) = Z. \tag{6.1.5}$$

We now introduce a G-submodule E of U by letting E be the subspace spanned by all elements of the form x^k, $k = 0, 1, \cdots$, where $x \in \mathfrak{g}$ is nilpotent.

Remark 30. Note that the universal enveloping algebra $U(\mathfrak{m})$ of $\mathfrak{m}$ is contained in E and (since every nilpotent element is conjugate to an element in $\mathfrak{m}$) that in fact E is the subspace of U spanned by all the algebras $\{U(\mathfrak{m}')\}$ where $\mathfrak{m}'$ runs through all the Lie subalgebras of $\mathfrak{g}$ conjugate to $\mathfrak{m}$.

THEOREM 20. *One has*

$$\delta(H)=E$$

where, we recall, H is the space of all G-harmonic polynomials on $\mathfrak{g}$.

Proof. Obviously $\delta(P')=P$ where P is the set of all nilpotent elements in $\mathfrak{g}$ and P' is defined as in § 1.4. But since $H=H_P$ by (4.8.3) the theorem follows immediately from (6.1.1). Q.E.D.

The filtration on U induces a filtration on E where $E_k=E\cap U_k$.

Now regard U as a Z-module (with respect to multiplication). The following is our main result on the structure of U.

THEOREM 21. *Let U be the universal enveloping algebra of a reductive Lie algebra* $\mathfrak{g}$. *Let Z be the center of U. Then U is free as a module over Z. In fact if E is the G-submodule of U defined above (see Remark 30) then the map*

$$Z\otimes E\rightarrow U \tag{6.1.6}$$

given by $p\otimes q\rightarrow pq$ is a G-module isomorphism.

Furthermore for any $\lambda\in D$ the multiplicity of the irreducible representation ν^λ in E is l_λ (the multiplicity of the zero weight for ν^λ). Moreover the order $o(\lambda)$ of λ (see § 5.3) is the smallest integer k such that the multiplicity of ν^λ is E_k is l_λ.

Proof. Let β denote the (G-module) map (6.1.6). To show first that β is surjective assume inductively that $U_j\subseteq \operatorname{Im}\beta$ (obviously $U_0\subseteq\operatorname{Im}\beta$ since $U_0\subseteq E$) for some integer j. Let $r\in U_{j+1}$. Then $r=\delta(g)$ where, by Theorem 11, § 4.8,

$$g=\sum_i f_ih_i$$

with $f_i\in J$, $h_i\in H$ and $\deg f_i+\deg h_i\leqq j+1$. But then

$$r=\beta(\sum_i\delta(f_i)\otimes\delta(h_i))\bmod U_j$$

by (6.1.2). We have, of course, used (6.1.5) and Theorem 20. But since $U_j\subseteq\operatorname{Im}\beta$ it follows that $r\in\operatorname{Im}\beta$. Hence β is surjective.

Now let $p\in Z\otimes E$ where $p\neq 0$. Then p is the image of an element $e\in J\otimes H$ under the isomorphism $J\otimes H\rightarrow Z\otimes E$ induced by δ (using (6.1.5) and Theorem 20). Furthermore we may assume that $e=\sum_i f_i\otimes h_i$ where $0\neq f_i\in J$ and the h_i are homogeneous and linearly independent in H.

Now let $k=\max_i(\deg f_i+\deg h_i)$. It follows therefore by Theorem 11, § 4.8, that if $g=\sum_i f_ih_i$ then $g\neq 0$ and $\deg g=k$. Hence

$$\delta(g)\neq 0 \bmod U_{k-1}. \tag{6.1.7}$$

But now $p = \sum_i \delta(f_i) \otimes \delta(h_i)$. Thus $\beta(p) = \sum_i \delta(f_i)\delta(h_i)$. But by (6.1.2) one has

$$\delta(g) = \beta(p) \bmod U_{k-1}. \tag{6.1.8}$$

Thus $\beta(p) \neq 0$ by (6.1.7-8) and hence β is an isomorphism. The remaining statements of the theorem follow immediately from Theorems 11, 17 and 20. Q. E. D.

6.2. Let G_1 be any algebraic reductive group whose Lie algebra is $\mathfrak{g}$. If

$$\nu\colon G_1 \to \operatorname{Aut} V \tag{6.2.1}$$

is a representation of G_1 on a finite dimensional space V then the corresponding representation of $\mathfrak{g}$ and U on V will also be denoted by ν.

An element $a \in G_1$ is called unipotent if a is of the form $a = \exp x$ where $x \in \mathfrak{g}$ is nilpotent.

LEMMA 16. *Let ν be as in* (6.2.1) *and let $W \subseteq \operatorname{End} V$ be the space spanned by all operators on V of the form $\nu(a)$ where $a \in G_1$ is unipotent. Then $W = \nu(E)$.*

Proof. By definition of E and the exponential formula it is obvious that $W \subseteq \nu(E)$. On the other hand if x is nilpotent then $\nu(\exp tx) \in W$ for all real numbers t. But, clearly, W also contains all t-derivatives of $\nu(\exp tx)$. Hence $\nu(x^k) \in W$ for $k = 0, 1, \cdots$. Thus $\nu(E) \subseteq W$. Q. E. D.

As a corollary of Theorem 21, § 6.1, we obtain

THEOREM 22. *Assume ν (as in* (6.2.1)*) is irreducible. Then*

$$\nu(E) = \operatorname{End} V \tag{6.2.2}$$

or equivalently (by Lemma 16*), every operator b on V may be put in the form*

$$b = \sum_{i=1}^{k} c_i \nu(a_i) \tag{6.2.3}$$

where the c_i are complex scalars and the a_i are unipotent elements of G_1.

Proof. Since ν is irreducible one has $\nu(U) = \operatorname{End} V$ and $\nu(Z) = \boldsymbol{C}$. The equation (6.2.2) then follows from the fact that (6.1.6) is an isomorphism. The second form of it follows from Lemma 16. Q. E. D.

Now let D_1, V_1^{ξ} and ν_1^{ξ}, for $\xi \in D_1$, play the same role for G_1 as the corresponding notation without the subscript 1 plays for G. (See § 2.1 and Remark 22, § 4.8).

Now $\operatorname{End} V_1^{\xi}$ is a G_1-module with respect to the tensor product of ν_1^{ξ} with the representation of G_1 contragradient to ν_1^{ξ}. But since the center of

G_1 obviously operates trivially on $\operatorname{End} V_1^{\xi}$ it follows that $\operatorname{End} V_1^{\xi}$ is a G-module and in fact the homomorphism of U onto $\operatorname{End} V_1^{\xi}$ induced by ν_1^{ξ} is a G-module epimorphism. Since all modules under consideration are completely reducible it follows therefore by Theorem 21, § 6.1, that ν_1^{ξ} induces a G-module epimorphism

(6.2.4) $$E^{\lambda} \to (\operatorname{End} V_1^{\xi})^{\lambda}$$

for every $\lambda \in D$.

Since every representation of G induces a representation of G_1 it is clear that $D \subseteq D_1$. One defines a partial ordering on D_1 by putting $\xi \geqq \xi'$, for $\xi, \xi' \in D_1$, whenever $\xi - \xi' \in D_1$. It is easy to see that D_1 is a directed set with respect to this ordering. ("Sufficiently large" will mean all $\xi \in D_1$ such that $\xi \geqq \xi'$ for some $\xi' \in D_1$.)

LEMMA 17. *For any $\xi \in D_1$ and $\lambda \in D$ let $n_{\lambda}(\xi)$ denote the multiplicity of ν^{λ} in* $\operatorname{End} V_1^{\xi}$ *(regarded as a G-module). Then*

(6.2.5) $$n_{\lambda}(\xi) \leqq l_{\lambda}$$

and (for fixed λ) the equality holds for ξ sufficiently large.

Proof. Since (6.2.4) is an epimorphism the inequality (6.2.5) is an immediate consequence of Theorem 21, § 6.1. However a much simpler and more direct proof of the inequality may be given using § 4.1 in [12].

Now identifying $V_1^{\xi} \otimes V_{1,\xi}$ with $\operatorname{End} V_1^{\xi}$ and regarding G-modules as G_1-modules it follows immediately from Schur's lemma, upon forming the triple tensor product $V_{\lambda} \otimes V_1^{\xi} \otimes V_{1,\xi}$, that $n_{\lambda}(\xi)$ is also the multiplicity of ν_1^{ξ} in $\nu_{\lambda} \otimes \nu_1^{\xi}$.

We refer now to [12], § 4.4, for the definition as to when λ is totally subordinate to ξ. By Theorem 5.1, (3) in [12] λ is totally subordinate to ξ for ξ sufficiently large. But now by (6) in this theorem (where $\mu = 0$, $\lambda_2 = \lambda$, $\lambda_1 = \xi$) the multiplicity of ν_1^{ξ} in $\nu_{\lambda} \otimes \nu_1^{\xi}$ is l_{λ} whenever λ is totally subordinate to ξ. Hence $n_{\lambda}(\xi) = l_{\lambda}$ whenever λ is totally subordinate to ξ or when ξ is sufficiently large. Q. E. D.

Harish-Chandra proved in [8] that if $Y \subseteq U$ is any one-dimensional subspace there exists $\xi \in D_1$ such that ν_1^{ξ} is faithful on Y. This is not true in general for higher dimensional subspaces. For example if $p \in Z$ and $q \in U$ where $q \neq 0$ and $p \notin U_0$ then q and pq span a two dimensional space in U but its image under ν_1^{ξ}, for any $\xi \in D_1$, is at most a one dimensional space. We now observe, however, that the generalization is true provided that $Y \subseteq E$.

THEOREM 23. *Let $Y \subseteq E$ be any finite dimensional subspace. Then the irreducible representation ν_1^{ξ} is faithful on Y for all $\xi \in D_1$ sufficiently large.*

Proof. Since Y is finite dimensional there exists k such that $Y \subseteq E_k$.

Now let $C \subseteq D$ denote the set of all $\lambda \in D$ such that ν^{λ} occurs with positive multiplicity in E_k. Since E is finite dimensional it is obvious that C is a finite set. Now since D_1 is a directed set it follows that equality holds in (6.2.5) for all $\lambda \in C$ and all ξ sufficiently large. But then by Theorem 21, § 6.1, the map (6.2.4) is an isomorphism also for all $\lambda \in C$ and $\xi \in D_1$ sufficiently large. Thus $\nu_1{}^{\xi}$ is faithful on E_k and hence on Y for all ξ sufficiently large. Q. E. D.

REFERENCES.

[1] A. Borel and Harish-Chandra, "Arithmetic subgroups of algebraic groups," *Annals of Mathematics*, vol. 75 (1962), pp. 485-535.

[2] C. Chevalley, "Invariants of finite groups generated by reflections," *American Journal of Mathematics*, vol. 77 (1955), pp. 778-782.

[3] ———, *Fondéments de la géométrie algébrique*, Course notes at the Institut Henri Poincaré, Paris, 1958.

[4] A. J. Coleman, "The Betti numbers of the simple Lie groups," *Canadian Journal of Mathematics*, vol. 10 (1958), pp. 349-356.

[5] E. B. Dynkin, "Semi-simple subalgebras of semi-simple Lie algebras," *American Mathematical Society Translations*, ser. 2, vol. 6 (1957), pp. 111-244.

[6] F. Gantmacher, "Canonical representation of automorphisms of a complex semi-simple Lie group," *Matematičeskiĭ Sbornik*, vol. 47 (1939), pp. 104-146.

[7] Harish-Chandra, "On a lemma of Bruhat," *Journal de Mathématiques Pures et Appliquées*, vol. 9, 315 (1956), pp. 203-210.

[8] ———, "On representations of Lie algebras," *Annals of Mathematics*, vol. 50 (1949), pp. 900-915.

[9] G. Hochschild and G. D. Mostow, "Representations and representative functions on Lie groups, III," *Annals of Mathematics* (2), vol. 70 (1959), pp. 85-100.

[10] S. Helgason, "Some results in invariant theory," *Bulletin of the American Mathematical Society*, vol. 68 (1962), pp. 367-371.

[11] N. Jacobson, "Completely reducible Lie algebras of linear transformations," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 105-133.

[12] B. Kostant, "A formula for the multiplicity of a weight," *Transactions of the American Mathematical Society*, vol. 9 (1959), pp. 53-73.

[13] ———, "The principal three-dimensional subgroup and the Betti numbers of a complex simple Lie group," *American Journal of Mathematics*, vol. 81 (1959), pp. 973-1032.

[14] M. Rosenlicht, "On quotient varieties and the affine embedding of certain homogeneous spaces," *Transactions of the American Mathematical Society*, vol. 101 (1961), pp. 211-223.

[15] J. P. Serre, "Faisceaux algébriques cohérent," *Annals of Mathematics*, vol. 61 1955), pp. 197-278.

[16] A. Seidenberg, "The hyperplane sections of normal varieties," *Transactions of the American Mathematical Society*, vol. 64 (1950), pp. 357-386.

[17] R. Steinberg, "Invariants of finite reflection groups," *Canadian Journal of Mathematics*, vol. 12 (1960), pp. 616-618.

[18] O. Zariski and P. Samuel, *Commutative Algebra*, vol. I, van Nostrand Company, Princeton, 1958.

[19] ———, *Commutative Algebra*, vol. II, van Nostrand Company, Princeton, 1960.

SOME DISTORTION THEOREMS OF QUASI-ANALYTIC MAPPINGS IN THE SPACE OF TWO COMPLEX VARIABLES.* [1]

By Stefan Bergman.

1. **Introduction.** In the theory of schlicht conformal mappings, often starting from some basic properties of these mappings and using additional considerations, we derive various distortion theorems.

Let $ds(z)$ denote the length of the line element, $da(z)$ the area element in the z-plane. If

$$z^* = z^*(z) \tag{1}$$

is a conformal transformation, then

$$\left(\frac{ds(z)}{ds(z^*)}\right)^2 = \left(\frac{da(z)}{da(z^*)}\right), \tag{2}$$

where $ds(z^*)$ and $da(z^*)$ are the corresponding quantities after the transformation.

Relation (2) can easily be generalized to the case of a large class of quasi-conformal transformations, namely the class of the mappings by a pair $(u(x,y), v(x,y))$ of differentiable functions which satisfy the generalized Cauchy-Riemann equations.

$$Cu_y + v_x = 0, \qquad Cu_x - v_y = 0, \qquad 0 < C < \infty, \tag{3}$$

where $C = C(x,y)$ is a continuously differentiable function. Mappings of this kind were introduced independently by Bers and Gelbart [B. G. 1], [B. G. 2] and by the author [B. 4]. They were extensively studied and generalized by Bers [B. 12] through [B. 16].

In the case of transformations by a pair of functions satisfying (3), it holds

$$\frac{\partial(u,v)}{\partial(x,y)} = C(u_x^2 + u_y^2) = \frac{1}{C}(v_x^2 + v_y^2) \tag{4}$$

from which the generalization of the relation (2) can be obtained. Using a

* Received March 15, 1963.

[1] This work was suported in part by the National Science Foundation Grant Number 21344 at Stanford University.

similar technique as in the case of conformal mappings, one obtains various distortion theorems for mappings by a pair of functions satisfying system (3).[2]

Generalizing the latter transformation, one introduces mappings of the domains of the (four-dimensional) x_1, y_1, x_2, y_2-space by four functions u_1, v_1, u_2, v_2 connected by system (2.1),[3] see p. 407. Mappings of this kind represent a generalization of pseudo-conformal transformations (PCT), they represent a subclass of quasi-pseudo-conformal transformations (QPCT).[4]

One shows that in the case of these mappings the relation (2.6), a generalization of (4), holds, and an analogue of (2) can be obtained if we replace the length (ds) of a line element by the so-called B-area of the surface element and the area element da by the volume element. Concerning the B-area compare Remark 1.1.

In the present paper we show that some of the consequences, analogous to those obtained by Ahlfors (for schlicht quasi-conformal mappings), can be made in the case of schlicht QPCT's. They follow from relations (2.6) and lead in this case to distortion theorems in the theory of QPCT's whose components satisfy relations (2.1).

This generalization is not immediate since the situation in the case of four-dimensional space is different from that of two-dimensional space. One of the difficulties which arises in these considerations is connected with the fact that the B-area of a surface has features different from those of the usual area, e.g., if a surface is analytic in the sense of the theory of two variables, then its B-area vanishes.

Remark 1.1. Let a surface $\mathfrak{S}^2$ be covered by a coordinate net. If we denote by ds_1, ds_2 the line elements taken in the direction of the coordinate net, then the element of $\mathfrak{S}^2$ equals $ds_1 ds_2 \sin\alpha$, where α is the angle between ds_1 and ds_2 The element of the B-area equals $ds_1 ds_2 \sin\beta$, where β is the *analytic* angle between ds_1 and ds_2 (see [B. 3], p. 476, and [B. 9]). If ds_k lies in the analytic plane $a^{(k)}z_1 + b^{(k)}z_2 = c_k$, then

$$(6) \qquad \sin\beta = \frac{|a^{(1)}b^{(2)} - a^{(2)}b^{(1)}|}{[(|a^{(1)}|^2 + |b^{(1)}|^2)(|a^{(2)}|^2 + |b^{(2)}|^2)]^{\frac{1}{2}}}$$

(see [B. 9], pp. 10-11).

Remark 1.2. In addition to the procedure based on the relation (2.6)

---

[2] Concerning the literature about distortion theorems in quasi-conformal mappings see, e.g., [K. 1].

[3] (2.1) = (1) of § 2.

[4] The theory of QPCT's is connected with the theory of the so-called functions of the extended class. Various types of functions of the extended class have been considered by Bremermann [B. 17], Lowdenslager [L. 3] and the author [B. 5].

and used in the present paper, one obtains for PCT's distortion theorems exploiting the connection between the kernel function and certain minimum problems, see [B. 2].

Remark 1.3. Using the theory of the kernel function, one can define an abstract manifold which is invariant with respect to PCT's (the so-called space of the class of pseudo-conformally equivalent domains (see [B. 6], p. 33)). The question how the invariants (curvature, etc.) of the basic space change if we apply a QPCT is studied in [B. 11].

Domains in the space of two complex variables have distinguished boundary sets, e. g., an analytic polyhedron has the distinguished boundary surface (see [B. 3]). In the case of PCT's regular at the boundary, these distinguished sets (invariant in QPCT's) will be discussed in a future publication.

Remark 1.4. Distortion theorems for mappings by solutions of systems of differential equations more general than (3) are discussed in [L. 1] and [L. 2].

2 A property of schlicht quasi-pseudo-conformal transformations. Let $C_\mu \equiv C_\mu(x_1, y_1, x_2, y_2)$, $\mu = 1, 2$, be continuously differentiable *positive* functions x_1, y_1, x_2, y_2 defined in a (bounded) domain $\mathfrak{D}$ of x_1, y_1, x_2, y_2-space. Further let u_k, v_k, $u_k \equiv u_k(x_1, y_1, x_2, y_2)$, $v_k \equiv v_k(x_1, y_1, x_2, y_2)$, $k = 1, 2$, $(x_1, y_1, x_2, y_2) \in D$, be four continuously differentiable functions which satisfy the system

$$(1) \qquad C_\mu \frac{\partial u_k}{\partial y_\mu} = -\frac{\partial v_k}{\partial x_\mu}, \qquad C_\mu \frac{\partial v_k}{\partial y_\mu} = \frac{\partial u_k}{\partial x_\mu}, \qquad k = 1, 2, \quad \mu = 1, 2.$$

A one-to-one mapping $\mathbf{W}$ of the domain $\mathfrak{D}$ by the functions u_1, v_1, u_2, v_2 satisfying the relation (1) is a QPCT of $\mathfrak{D}$.

Let $\mathfrak{S}^2$, $\mathfrak{S}^2 \subset \mathfrak{D}$, be a segment of the surface

$$(2) \qquad x_1 = x_1^* = \text{const.} \qquad x_2 = x_2^* = \text{const.}$$

and let $\mathfrak{T}^2 = \mathbf{W}(\mathfrak{S}^2)$ be the image of $\mathfrak{S}^2$ in the u_1, v_1, u_2, v_2-space.

In addition to the area of a surface element we can consider the so-called B-area [5] of the element of $\mathfrak{S}^2$, given by

$$(3) \qquad d\mathrm{B}(\mathfrak{S}^2) = dy_1 dy_2.$$

[5] Concerning the notion of the B-area see [B. 3], pp. 475 ff.

THEOREM 1.[6] *The* B-*area* $dB(\mathfrak{T}^2)$ *of the surface element of* $\mathfrak{T}^2 - \mathbf{W}(\mathfrak{S}^2)$ *is given by*

$$dB(\mathfrak{T}^2) = \left[C_1^{-1} C_2^{-1} \left(\frac{\partial(u_1, v_1, u_2, v_2)}{\partial(x_1, y_1, x_2, y_2)} \right) \right]^{\frac{1}{2}} dy_1 dy_2. \tag{4}$$

Proof. According to the definition (see [B. 3], p. 475)

$$dB(\mathfrak{T}^2) = \left| \frac{\partial(\tilde{w}_1, \tilde{w}_2)}{\partial(y_1, y_2)} \right| dy_1 dy_2 \tag{5}$$

$$\tilde{w}_k \equiv w_k(z_1, z_2) \equiv u_k(x_1, y_1, x_2, y_2) + i v_k(x_1, y_1, x_2, y_2).$$

We shall show that

$$\left| \frac{\partial(\tilde{w}_1, \tilde{w}_2)}{\partial(y_1, y_2)} \right| = \left(C_1^{-1} C_2^{-1} \frac{\partial(u_1, v_1, u_2, v_2)}{\partial(x_1, y_1, x_2, y_2)} \right)^{\frac{1}{2}}. \tag{6}$$

A formal computation shows that

$$\begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial v_1}{\partial x_1} & \frac{\partial u_2}{\partial x_1} & \frac{\partial v_2}{\partial x_1} \\ \frac{\partial u_1}{\partial y_1} & \frac{\partial v_1}{\partial y_1} & \frac{\partial u_2}{\partial y_1} & \frac{\partial v_2}{\partial y_1} \\ \frac{\partial u_1}{\partial x_2} & \frac{\partial v_1}{\partial x_2} & \frac{\partial u_2}{\partial x_2} & \frac{\partial v_2}{\partial x_2} \\ \frac{\partial u_1}{\partial y_2} & \frac{\partial v_1}{\partial y_2} & \frac{\partial u_2}{\partial y_2} & \frac{\partial v_2}{\partial y_2} \end{vmatrix} = C_1 C_2 \left\{ -2 \frac{\partial(u_1, v_2)}{\partial(y_1, y_2)} \frac{\partial(u_2, v_2)}{\partial(y_1, y_2)} + \left[\frac{\partial(v_1, v_2)}{\partial(y_1, y_2)} \right]^2 + \left[\frac{\partial(v_1, u_2)}{\partial(y_1, y_2)} \right]^2 + \left[\frac{\partial(u_1, v_2)}{\partial(y_1, y_2)} \right]^2 + \left[\frac{\partial(u_1, u_2)}{\partial(y_1, y_2)} \right]^2 \right\}. \tag{7}$$

On the other hand,

$$\begin{aligned} \left| \frac{\partial(u_1 + iv_1, u_2 + iv_2)}{\partial(y_1, y_2)} \right|^2 = \left[\frac{\partial(u_1, u_2)}{\partial(y_1, y_2)} \right]^2 + \left[\frac{\partial(v_1, v_2)}{\partial(y_1, y_2)} \right]^2 + \left[\frac{\partial(u_1, v_2)}{\partial(y_1, y_2)} \right]^2 \\ + \left[\frac{\partial(v_1, u_2)}{\partial(y_1, y_2)} \right]^2 - 2 \frac{\partial(u_1, u_2)}{\partial(y_1, y_2)} \frac{\partial(v_1, v_2)}{\partial(y_1, y_2)} \\ + 2 \frac{\partial(u_1, v_2)}{\partial(y_1, y_2)} \frac{\partial(v_1, u_2)}{\partial(y_1, y_2)}. \end{aligned} \tag{8}$$

and

$$\frac{\partial(u_1, u_2)}{\partial(y_1, y_2)} \frac{\partial(v_1, v_2)}{\partial(y_1, y_2)} - \frac{\partial(u_1, v_2)}{\partial(y_1, y_2)} \frac{\partial(v_1, u_2)}{\partial(y_1, y_2)} = \frac{\partial(u_1, v_1)}{\partial(y_1, y_2)} \cdot \frac{\partial(u_2, v_2)}{\partial(y_1, y_2)} \tag{9}$$

(6) follows from (7), (8) and (9).

[6] See also [B. 10], p. 137.

3. A generalization of an Ahlfors result to the case of two complex variables. In the case of one variable, Ahlfors [A. 1], [N. 1], p. 87, considers a curved linear strip $\mathfrak{B}^2$ in the x, y-plane. This strip is mapped by the function $w(z) = u(x,y) + iv(x,y)$ conformally and schlicht onto a strip bounded by lines $v = +a/2$ and $v = -a/2$ in the u, v-plane. He then obtains the inequality.

$$(1) \qquad a^2 \int_{x^{(1)}}^{x^{(2)}} \frac{dx}{l[\mathfrak{Q}^1(x)]} + \int_{x^{(1)}}^{x^{(2)}} \frac{\omega^2(x)\,dx}{l[\mathfrak{Q}^1(x)]} \leqq A \leqq a[u_2(x_2) - u_1(x_1)].$$

Here A is the area of the image of $\mathfrak{B}^2$ in the u, v-plane (see for details [A. 1] or [N. 1], p. 88 ff.),

$$u_1(x) = \min_{y \in \mathfrak{Q}^1(x)} u(x,y), \qquad u_2(x) = \max_{y \in \mathfrak{Q}^1(x)} u(x,y),$$
$$\omega(x) = u_2(x) - u_1(x),$$

$(x, 0) \in \mathfrak{Q}^1(x)$, $\mathfrak{Q}^1(x^*)$, is the connected part of the intersection of the strip with the line $x = x^*$, $l(\mathfrak{Q}^1(x^*)) =$ the length of $\mathfrak{Q}^1(x^*)$. In the present paper we generalize Ahlfors considerations to the case of QPCT's, whose components satisfy (2.1).

We consider in the x_1, y_1, x_2, y_2-space the domain $\mathfrak{D}$ bounded by segments of four hypersurfaces

$$(2a) \qquad y_1 = \tilde{y}_1(x_1, x_2, y_2), \qquad a_3 \leqq y_2 \leqq a_4, \qquad a_n \equiv a_n(x_1, x_2),$$
$$(2b) \qquad y_1 = \bar{\tilde{y}}_1(x_1, x_2, y_2), \qquad a_3^* \leqq y_2 \leqq a_4^*, \qquad a_n^* \equiv a_n^*(x_1, x_2),$$
$$n = 3, 4, 5, 6,$$

$$(3a) \qquad y_2 = \tilde{y}_2(x_1, x_2, y_1), \qquad a_5 \leqq y_1 \leqq a_6,$$
$$(3b) \qquad y_2 = \bar{\tilde{y}}_2(x_1, x_2, y_1), \qquad a_5^* \leqq y_1 \leqq a_6^*.$$

$\tilde{y}_k(x_1, x_2, y_{3-k})$, $\bar{\tilde{y}}_k(x_1, x_2, y_{3-k})$, $k = 1, 2$, are assumed to be single-valued continuously differtiable functions of (x_1, x_2, y_{3-k}) and the segments (2a), (2b), as well as (3a), (3b), do not intersect each other.

Remark 3.1. Here a_3, a_3^*, a_4, a_4^*, a_5, a_5^*, a_6, a_6^* are conveniently chosen quantities. We note that these quantities are connected by the relations

$$a_3 = \tilde{y}_2(a_5), \qquad a_4 = \bar{\tilde{y}}_2(a_5^*), \qquad \tilde{y}_2(a_5) \equiv \tilde{y}_2(a_5, x_1, x_2)$$
$$a_5 = \tilde{y}_1(a_3), \qquad a_5^* = \tilde{y}_1(a_4),$$
$$a_4^* = \bar{\tilde{y}}_2(a_6^*), \qquad a_6 = \bar{\tilde{y}}_1(a_3^*),$$
$$a_6^* = \bar{\tilde{y}}_1(a_4^*), \qquad a_3^* = \tilde{y}_2(a_6^*)$$

We construct in the u_1, v_1, u_2, v_2-space the domain $\mathfrak{A}^*$ whose boundary consists of segments of hypersurfaces

(4a) $\quad v_1 = -a_1/2 \qquad$ (5a) $\quad v_2 = -a_2/2$

(4b) $\quad v_1 = a_1/2 \qquad$ (5b) $\quad v_2 = a_2/2$

and we assume that the schlicht (one-to-one) QPCT

$$\mathbf{W}: w_1 = w_1(z_1, z_2), \qquad w_2 = w_2(z_1, z_2) \tag{6}$$

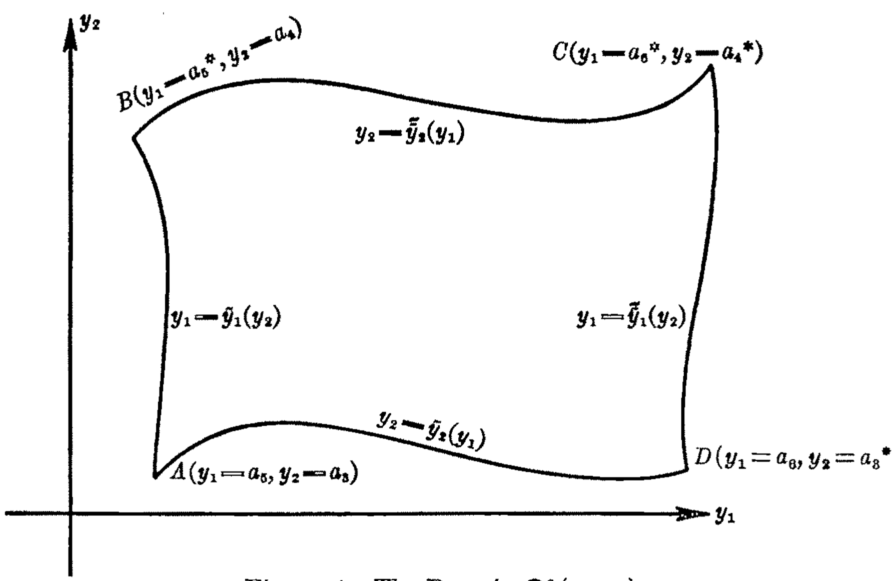


Figure 1. The Domain $\mathfrak{Q}^2(x_1, x_2)$.

maps the domain $\mathfrak{D}$ onto a domain $\mathfrak{A}$ which is a part of $\mathfrak{A}^*$. Let $\mathfrak{Q}^2(x_1^*, x_2^*)$ be a connected segment representing the part of

$$\mathfrak{S}^2(x_1^*, x_2^*) = \bar{\mathfrak{D}} \cap (x_1 = x_1^*, x_2 = x_2^*) \tag{7}$$

which includes the points $y_1 = 0, y_2 = 0$ in its interior. For simplicity's sake we assume that $\mathfrak{Q}^2(x_1, x_2) = \mathfrak{S}^2(x_1, x_2)$, $x_k^{(1)} \leqq x_k \leqq x_k^{(2)}$, $k = 1, 2$. $\mathfrak{Q}^2(x_1^*, x_2^*)$ is bounded by four curvilinear line segments $\mathfrak{t}_1^1, \mathfrak{t}_2^1, \mathfrak{t}_3^1, \mathfrak{t}_4^1$. Here

(8a) $\quad \mathfrak{t}_1^1(x_1^*, x_2^*) = [x_1 = x_1^*, x_2 = x_2^*, y_1 = \bar{y}_1(x_1^*, x_2^*, y_2), y_2 \in \mathfrak{p}_1^1(x_1, x_2)]$, $\mathfrak{p}_1^1(x_1^*, x_2^*)$ is the projection of $\mathfrak{t}_1^1(x_1^*, x_2^*)$ on the segment $a_3 \leqq y_2 \leqq a_4$ of the y_2-axis. Analogously,

(8b) $t_3^1(x_1^*, x_2^*) = [x_1 = x_1^*, x_2 = x_2^*, y_1 = \tilde{y}_1(x_1^*, x_2^*, y_2), y_2 \in \mathfrak{p}_3^1(x_1^*, x_2^*)]$

(8c) $\mathfrak{t}_2^1(x_1^*, x_2^*) = [x_1 = x_1^*, x_2 = x_2^*, y_2 = \tilde{y}_2(x_1^*, x_2^*, y_1), y_1 \in \mathfrak{p}_2^1(x_1^*, x_2^*)]$

(8d) $\mathfrak{t}_4^1(x_1^*, x_2^*) = [x_1 = x_1^*, x_2 = x_2^*, y_2 = \bar{y}_2(x_1^*, x_2^*, y_1), y_1 \in \mathfrak{p}_4^1(x_1^*, x_2^*)]$.

We assume that the domain $\mathfrak{D}$ and the functions w_1, w_2 are chosen so that

$$
(9) \qquad
\begin{aligned}
&\mathrm{Im}[w_1(x_1^*, x_2^*, \tilde{y}_1(y_2), y_2)] = -a_1/2 = \text{const.}\\
&\mathrm{Im}[w_1(x_1^*, x_2^*, \bar{y}_1(y_2), y_2)] = a_1/2 = \text{const.}\\
&\mathrm{Im}[w_2(x_1^*, x_2^*, y_1, \tilde{y}_2(y_1))] = -a_2/2 = \text{const.}\\
&\mathrm{Im}[w_2(x_1^*, x_2^*, y_1, \bar{y}_2(y_1))] = a_2/2 = \text{const.}
\end{aligned}
$$

The image $\mathbf{W}(\mathfrak{Q}^2(x_1, x_2))$ of $\mathfrak{Q}^2(x_1, x_2)$ in the w_1, w_2-space will be called $\mathfrak{L}^2(x_1, x_2) \equiv \tilde{\mathfrak{L}}^2(u_1, v_1, u_2, v_2)$, its boundary (the image of $\bigcup_{\nu=1}^{4} \mathfrak{t}_\nu^1(x_1, x_2)$) will be called $\mathfrak{l}^1(x_1, x_2)$.

The one-parameter family

$$(10) \qquad \mathfrak{h}^3(x_1) = \bigcup_{x_2^{(1)} \leqq \xi_2 \leqq x_2^{(2)}} \mathfrak{L}^2(x_1, \xi_2)$$

is a segment of a hypersurface. Its boundary is

$$(11) \qquad \mathfrak{H}^2(x_1) = \bigcup_{x_2^{(1)} \leqq \xi_2 \leqq x_2^{(2)}} \mathfrak{l}^1(x_1, \xi_2).$$

Since w_1, w_2 are continuous functions,

$$(12) \qquad \mathfrak{A} = \bigcup_{x_1^{(1)} \leqq x_1 \leqq x_1^{(2)}} \mathfrak{h}^3(x_1) = \bigcup_{\substack{x_1^{(1)} \leqq x_1 \leqq x_1^{(2)} \\ x_2^{(1)} \leqq x_2 \leqq x_2^{(2)}}} \bigcup \mathfrak{L}^2(x_1, x_2)$$

Since the mapping $\mathbf{W}$ is one-to-one,

$$(13) \qquad \mathfrak{L}^2(x_1, x_2) \cap \mathfrak{L}^2(x_1^*, x_2^*) = \emptyset \quad \text{if} \quad (x_1, x_2) \neq (x_1^*, x_2^*).$$

Indeed, if the left-hand side of (13) were not empty to the intersection point of $\mathfrak{L}^2(x_1, x_2)$ and $\mathfrak{L}^2(x_1^*, x_2^*)$ would correspond to two different points (x_1, x_2) and (x_1^*, x_2^*), this is impossible since the mapping $\mathbf{W}$ is one-to-one.

Let

$$(14) \qquad b(u_1, v_1, u_2, v_2) = \frac{d\mathrm{B}[\tilde{\mathfrak{L}}^2(u_1, v_1, u_2, v_2)]}{d\mathrm{A}[\tilde{\mathfrak{L}}^2(u_1, v_1, u_2, v_2)]}$$

be the ratio between the B- and the ordinary area of the surface element of $\mathfrak{L}^2(x_1, x_2)$.

Further, let $dp_\nu(u_1, v_1, u_2, v_2)$, $\nu = 1, 2, 3, 4, 5, 6$, the the Euclidean areas of the projection of the surface element of $\mathfrak{L}^2(x_1, x_2) = \mathbf{W}(\mathfrak{Q}^2(x_1, x_2))$

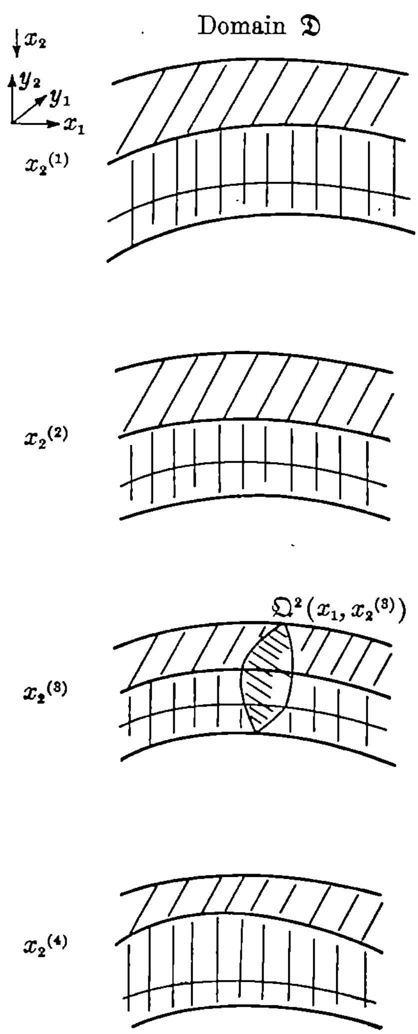


Fig. 2. Intersections of $\mathfrak{D}$ with $x_2 = x_2^{(\nu)} =$ const.

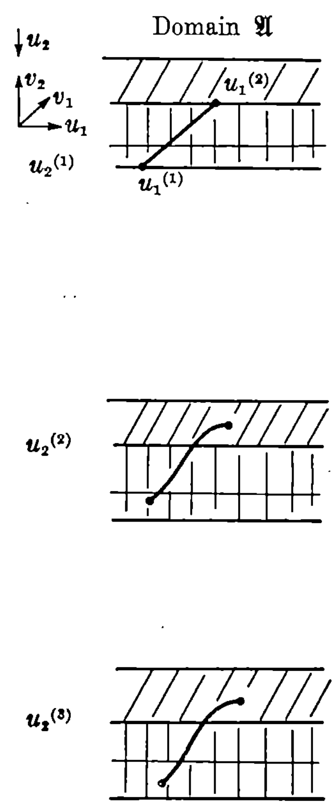


Fig. 3. Intersections of $\mathfrak{A}$ with $u_2 = u_2^{(\nu)} =$ const.

on the u_1, v_1-, u_1, u_2-, u_1, v_2-, v_1, u_2-, v_1, v_2- and u_2, v_2-planes, respectively. $P_\nu[\mathfrak{L}^2(x_1, x_2)]$, $\nu = 1, 2, 3, 4, 5, 6$, are the projection of $\mathfrak{L}^2(x_1, x_2)$ on the before mentioned planes. We note that

$$[dA[\mathfrak{L}^2(u_1, v_1, u_2, v_2)]]^2 = \sum_{\nu=1}^{6} [dp_\nu(u_1, v_1, u_2, v_2)]^2,$$

$$u_k + iv_k = u_k(z_1, z_2) + iv_k(z_1, z_2), \quad (u_1, v_1, u_2, v_2) \in \mathfrak{L}^2(x_1, x_2). \tag{15}$$

THEOREM 2. *Let* $\mathbf{W}$ *be a schlicht* QPCT *of* $\mathfrak{D}$ *onto the domain* $\mathfrak{A}$, *satisfying the conditions* (9). *Then*

$$\frac{1}{a_1 a_2} \sum_{\nu=1}^{6} \int_{x_2^{(1)}}^{x_2^{(2)}} \int_{x_1^{(1)}}^{x_1^{(2)}} \frac{[b_m(x_1, x_2) \mathfrak{P}_\nu[\mathfrak{L}^2(x_1, x_2)]]^2}{A[\mathfrak{Q}^2(x_1, x_2)]} dx_1 dx_2$$

$$\leqq \prod_{k=1}^{2} (1/C_k^{(\min)}) [m_k^{(2)} - m_k^{(1)}], \tag{16}$$

$$C_k^{(\min)} = \min_{(x_1, y_1, x_2, y_2) \in \mathfrak{D}} C_k(x_1, y_1, x_2, y_2). \tag{16a}$$

Here

$$\begin{aligned} m_1^{(2)} &= \max[u_1(x_1, y_1, x_2, y_2)], (y_1, y_2) \in \mathfrak{Q}^2(x_1^{(2)}, \zeta_2), x_2^{(1)} \leqq \zeta_2 \leqq x_2^{(2)} \\ m_1^{(1)} &= \min[u_1(x_1, y_1, x_2, y\ )], (y_1, y_2) \in \mathfrak{Q}^2(x_1^{(1)}, \zeta_2), x_2^{(1)} \leqq \zeta_2 \leqq x_2^{(2)} \end{aligned} \tag{17}$$

$$\begin{aligned} m_2^{(2)} &= \max[u_2(x_1, y_1, x_2, y_2)], (y_1, y_2) \in \mathfrak{Q}^2(\zeta_1, x_2^{(2)}), x_1^{(1)} \leqq \zeta_1 \leqq x_1^{(2)} \\ m_2^{(1)} &= \min[u_2(x_1, y_1, x_2, y_2)], (y_1, y_2) \in \mathfrak{Q}^2(\zeta_1, x_2^{(1)}), x_1^{(1)} \leqq \zeta_1 \leqq x_1^{(2)} \end{aligned} \tag{18}$$

$$b_m(x_1, x_2) = \min[b(u_1, v_1, u_2, v_2)], \qquad (u_1, v_1, u_2, v_2) \in \mathfrak{L}^2(x_1, x_2). \tag{19}$$

Proof. By the Schwarz inequality

$$\left[\iint_{\mathfrak{Q}^2(x_1, x_2)} \left|\frac{\partial(w_1, w_2)}{\partial(y_1, y_2)}\right| dy_1 dy_2\right]^2 \leqq \iint_{\mathfrak{Q}^2(x_1, x_2)} dy_1 dy_2 \iint_{\mathfrak{Q}^2(x_1, x_2)} \left|\frac{\partial(w_1, w_2)}{\partial(y_1, y_2)}\right|^2 dy_1 dy_2 \tag{20}$$

or

$$\frac{[B[\mathfrak{L}^2(x_1, x_2)]]^2}{A[\mathfrak{Q}^2(x_1, x_2)]} \leqq \iint_{\mathfrak{Q}^2(x_1, x_2)} \left|\frac{\partial(w_1, w_2)}{\partial(y_1, y_2)}\right|^2 dy_1 dy_2. \tag{21}$$

Multiplying both sides of (21) by $dx_1 dx_2$, integrating over $[x_k^{(1)} \leqq x_k \leqq x_k^{(2)}, k = 1, 2]$ and using (2.6) and (16a), we obtain

$$\int_{x_1^{(1)}}^{x_1^{(2)}} \int_{x_2^{(1)}}^{x_2^{(2)}} \frac{[B[\mathfrak{L}^2(x_1, x_2)]]^2}{A[\mathfrak{Q}^2(x_1, x_2)]} dx_1 dx_2 \leqq (1/C_1^{(\min)} C_2^{(\min)}) V(\mathfrak{A}) \tag{22}$$

where $V(\mathfrak{A})$ is the volume of the domain covered by $\mathfrak{L}^2(x_1, x_2)$, $x_k^{(1)} \leqq x_k \leqq x_k^{(2)}$, $k=1,2$. In accordance with (17), (18), since the mapping is schlicht,

$$V(\mathfrak{A}) \leqq \prod_{k=1}^{2} a_k (m_k^{(2)} - m_k^{(1)}) \tag{23}$$

According to (14) and (15),

$$\begin{aligned} dB(\mathfrak{L}^2(x_1, x_2)) &= b(u_1, v_1, u_2, v_2)\, dA(\mathfrak{L}^2(x_1, x_2)) \\ &\geqq \frac{1}{\sqrt{6}} b(u_1 v_1, u_2, v_2) \sum_{\nu=1}^{6} dp_\nu(u_1, v_1, u_2, v_2) \end{aligned} \tag{24}$$

$$u_k = u_k(z_1, z_2),\ v_k = v_k(z_1, z_2),\ k=1,2,\ (u_1, v_1, u_2, v_2) \in \mathfrak{L}^2(x_1, x_2).$$

We note from (15) it follows that

$$dA(u_1, v_1, u_2, v_2) \geqq \frac{1}{\sqrt{6}} \sum_{\nu=1}^{6} b(u_1, v_1, u_2, v_2)\, dp_\nu(u_1, v_1, u\ , v_2). \tag{25}$$

From (22), (23), (24) follows that the left-hand side of (22) is larger than

$$\begin{aligned} &\int_{x_1^{(1)}}^{x_1^{(2)}} \int_{x_2^{(1)}}^{x_2^{(2)}} \sum_{\nu=1}^{6} \frac{\left[\displaystyle\iint_{(u_1, v_1, u_2, v_2) \in \mathfrak{L}^2(x_1, x_2)} b(u_1, v_1, u_2, v_2)\, dp_\nu(u_1, v_1, u_2, v_1) \right]^2}{A[\mathfrak{L}^2(x_1, x_2)]} dx_1 dx_2 \\ &\geqq \int_{x_1^{(1)}}^{x_1^{(2)}} \int_{x_2^{(1)}}^{x_2^{(2)}} \sum_{\nu=1}^{6} \left\{ \frac{b_{\min}(x_1, x_2)\, A[\mathfrak{P}_\nu(\mathfrak{L}^2(x_1, x_2))]^2}{A[\mathfrak{L}^2(x_1, x_2)]} \right\} dx_1 dx_2. \end{aligned} \tag{26}$$

(16) follows from (22), (23) and 26).

We make now the hypothesis that the function $b_m(x_1, x_2)$, see (19), has a positive lower bound, namely

$$0 < \beta \leqq b_m(x_1, x_2). \tag{27}$$

Let

$$\begin{aligned} &\mathfrak{P}^2_{u_1 v_1}(x_1, x_2) = P_1[\mathfrak{L}^2(x_1, x_2)] \equiv \mathfrak{P}_1^2, \mathfrak{P}^2_{u_1 v_2}(x_1, x_2) = P_2[\mathfrak{L}^2(x_1, x_2)] \equiv \mathfrak{P}_2^2 \\ &\mathfrak{P}^2_{u_2 v_1}(x_1, x_2) = P_3[\mathfrak{L}^2(x_1, x_2)] \equiv \mathfrak{P}_3^2, \mathfrak{P}^2_{u_2 v_2}(x_1, x_2) = P_4[\mathfrak{L}^2(x_1, x_2)] \equiv \mathfrak{P}_4^2 \\ &\mathfrak{P}^2_{v_1 v_2}(x_1, x_2) = P_5[\mathfrak{L}^2(x_1, x_2)] \equiv \mathfrak{P}_5^2, \mathfrak{P}^2_{u_1 u_2}(x_1, x_2) = P_6[\mathfrak{L}^2(x_1, x_2)] \equiv \mathfrak{P}_6^2 \end{aligned} \tag{28}$$

be the projections of $\mathfrak{L}^2(x_1, x_2)$ on the corresponding plane.

Let

$$\omega_k(x_1, x_2) \equiv \max_{(y_1, y_2) \in \mathfrak{L}^2(x_1, x_2)} [u_k(x_1, y_1, x_2, y_2)] - \min_{(y_1, y_2) \in \mathfrak{L}^2(x_1, x_2)} [u_k(x_1, y_1, x_2, y_2)]. \tag{29}$$

LEMMA. *$\omega_k(x_1, x_2)$ is the length of the projection of $\mathfrak{L}^2(x_1, x_2)$ on the u_k-axis.*

Proof. Suppose that a point $u_k^{(0)}$

$$(30) \qquad \min_{(y_1,y_2)\in\mathfrak{D}^2(x_1,x_2)} [u_k(x_1,y_1,x_2,y_2)] < u_k^{(0)} < \max_{(y_1,y_2)\in\mathfrak{D}^2(x_1,x_2)} [u_k(x_1,y_1,x_2,y_2)]$$

does not belong to the projection of $\mathfrak{L}^2(x_1,x_2)$ on the u_k-axis. This would mean that the hyperplane $u_k = u_k^{(0)}$ would divide $\mathfrak{L}^2(x_1,x_2)$ into two disconnected parts. This is impossible since $\mathfrak{D}^2(x_1,x_2)$ is connected and the mapping $\mathbf{W}$ is continuous and one-to-one.

Notation. $\Omega(x_1,x_2)$ is the area of the projection of $\mathfrak{L}^2(x_1,x_2)$ on the u_1,u_2-plane.

THEOREM 2. *If the* $b_m(x_1,x_2)$ *satisfy the condition* (27), *then* (16) *in Theorem* 1 *can be replaced by*

$$(31) \qquad \frac{\beta^2}{a_1a_2}\int_{x_1^{(1)}}^{x_1^{(2)}}\int_{x_2^{(1)}}^{x_2^{(2)}} \frac{[(a_1^2+a_2^2)(\omega_1^2(x_1,x\ )+\omega_2^2(x_1,x_2))+a_1^2a_2^2+\Omega^2(x_1,x_2)]}{A[\mathfrak{D}^2(x_1,x_2)]}\,dx_1dx_2$$
$$\leqq \prod_{k=1}^{2} (1/C_k^{(\min)})\,[m_k^{(2)}-m_k^{(1)}].$$

Proof. According to the lemma and (28)

$$(32) \qquad \begin{aligned} &A[\mathfrak{P}^2_{u_sv_k}(x_1,x_2)] = a_k\omega_s(x_1,x_2), \quad k=1,2, \quad s=1,2,\\ &A[\mathfrak{P}^2_{u_1u_2}(x_1,x_2)] = \Omega(x_1,x_2), \quad A[\mathfrak{P}^2_{v_1v_2}(x_1,x_2)] = a_1a_2, \end{aligned}$$

and therefore

$$(33) \qquad \begin{aligned} &\sum_{\nu=1}^{6}\int_{x_1^{(1)}}^{x_1^{(2)}}\int_{x_2^{(1)}}^{x_2^{(2)}} \frac{\{b_m(x_1,x_2)[A(\mathfrak{P}_\nu^2(x_1,x_2))]\}^2}{A[\mathfrak{D}^2(x_1,x_2)]}\,dx_1dx_2\\ &\geqq \beta\int_{x_1^{(1)}}^{x_1^{(2)}}\int_{x_2^{(1)}}^{x_2^{(2)}} \frac{(a_1^2+a_2^2)(\omega^2(x_1,x_2)+\omega_2^2(x_1,x_2))+a_1^2a_2^2+\Omega^2(x_1,x_2)}{A[\mathfrak{D}^2(x_1,x_2)]}\,dx_1dx_2 \end{aligned}$$

Our considerations yield further bounds for the sum of the squares of some projection areas. These bounds do *not* depend on the function b, see (14), but they involve the quantity $\delta^{(\max)}$ to be described in the following.

In every $\mathfrak{D}^2(x_1,x_2)$, x_1, x_2 fixed, $u_k(x_1,y_1,x_2,y_2)$, $v_k(x_1,y_1,x_2,y_2)$, $k=1,2$ are functions of y_1, y_2. The pair u_k, v_k, $k=1$ or 2 defines a mapping $\mathbf{M}_k$ of the segment $\mathfrak{D}^2(x_1,x_2)$ on a domain of the u_k,v_k-plane. Since the functions u_k, v_k are continuously differentiable in $\mathfrak{D}^2(x_1,x_2)$,

$$(34) \qquad \left|\frac{\partial(u_k,v_k)}{\partial(y_1,y_2)}\right| = |D_k(x_1,y_1,x_2,y_2)|$$

is uniformly bounded in (closed) $\mathfrak{Q}^2(x_1, x_2)$, $x_k^{(1)} \leqq x_k \leqq x_k^{(2)}$, $k = 1, 2$. ($|D_k(y_1, y_2)|$ is the distortion of the area element at the point y_1, y_2 of $\mathfrak{Q}^2(x_1, x_2)$ when applying the transformation $\mathbf{M}_k$.)

Let

$$(35) \qquad \delta^{(\max)} = \max \left| \prod_{k=1}^{2} D_k(x_1, y_1, x_2, y_2) \right| < \infty .$$

Here the maximum is taken for

$$(y_1, y_2) \in \mathfrak{Q}^2(x_1, x_2), \quad x_k^{(1)} \leqq x_k \leqq x_k^{(2)}.$$

THEOREM 3. *Let the* QPCT (6), *satisfying the conditions* (9), *map* $\mathfrak{D}$ *onto the domain* $\mathfrak{A}$. *Then for the projections* $\Omega(x_1, x_2)$, $\omega_k(x_1, x_2)$ *of* $\mathfrak{Q}^2(x_1, x_2)$, $x_k^{(1)} \leqq x_k \leqq x_k^{(2)}$, $k = 1, 2$, *the inequality*

$$(36) \qquad \int_{x_1^{(1)}}^{x_1^{(2)}} \int_{x_2^{(1)}}^{x_2^{(2)}} \frac{\Omega^2(x_1, x_2) + a_1^2 \omega_2^2(x_1, x_2) + a_2^2 \omega_1^2(x_1, x_2) + a_1^2 a_2^2}{\mathrm{A}[\mathfrak{Q}^2(x_1, x)]} dx_1 dx_2 \leqq \left(\prod_{k=1}^{2} \frac{1}{C_k^{(\min)}} \right) \mathrm{V}(\mathfrak{A}) + 2\delta^{(\max)} \mathrm{V}(\mathfrak{D})$$

holds. $\mathrm{V}(\cdots) =$ *volume of* $\cdots$.

Proof. By (2.8) and (2.9)

$$(37) \qquad \left[\frac{\partial(u_1, v_2)}{\partial(y_1, y_2)}\right]^2 + \left[\frac{\partial(u_1, v_2)}{\partial(y_1, y_2)}\right]^2 + \left[\frac{\partial(v_1, u_2)}{\partial(y_1, y_2)}\right]^2 + \left[\frac{\partial(v_1, v_2)}{\partial(y_1, y_2)}\right]^2 \leqq \left| \frac{\partial(w_1, w_2)}{\partial(y_1, y_2)} \right|^2 + 2 \prod_{k=1}^{2} \left| \frac{\partial(u_k, v_k)}{\partial(y_1, y_2)} \right| .$$

Multiplying (37) by $dy_1 dy_2$, integrating over $\mathfrak{D}^2(x_1, x_2)$ and using (37), (34), (35), we obtain

$$(38) \qquad \begin{aligned} &\iint_{\mathfrak{Q}^2(x_1,x_2)} \left[\frac{\partial(u_1, u_2)}{\partial(y_1, y_2)}\right]^2 dy_1 dy_2 + \iint_{\mathfrak{Q}^2(x_1,x_2)} \left[\frac{\partial(v_1, u_2)}{\partial(y_1, y_2)}\right]^2 dy_1 dy_2 \\ &+ \iint_{\mathfrak{Q}^2(x_1,x_2)} \left[\frac{\partial(v_1, v_2)}{\partial(y_1, y_2)}\right]^2 dy_1 dy_2 + \iint_{\mathfrak{Q}^2(x_1,x_2)} \left[\frac{\partial(u_1, u_2)}{\partial(y_1, y_2)}\right]^2 dy_1 dy_2 \\ &\leqq \iint_{\mathfrak{Q}^2(x_1,x_2)} \left| \frac{\partial(w_1, w_2)}{\partial(y_1, y_2)} \right|^2 dy_1 dy_2 + 2\delta^{(\max)} \iint_{\mathfrak{Q}^2(x_1,x_2)} dy_1 dy_2 . \end{aligned}$$

Applying the Schwarz inequality to the left hand side of (38) and using the notation introduced in (28), we have

$$(39)\quad \frac{\{A[P_{u_1u_2}(\mathfrak{L}^2)]\}^2+\{A[P_{u_1v_2}(\mathfrak{L}^2)]\}^2+\{A[P_{v_1u_2}(\mathfrak{L}^2)]\}^2+\{A[P_{v_1v_2}(\mathfrak{L}^2)]\}^2}{A[\mathfrak{Q}^2(x_1,x_2)]}$$
$$\leqq \iint_{\mathfrak{Q}^2(x_1,x_2)}\left|\frac{\partial(w_1,w_2)}{\partial(y_1,y_2)}\right|^2 dy_1dy_2+2\delta^{(\max)}\iint_{\mathfrak{Q}^2(x_1,x_2)} dy_1dy_2$$

Multiplying (39) by dx_1dx_2, integrating over $[x_k^{(1)}\leqq x_k\leqq x_k^{(2)},\ k=1,2]$, using (2.6) and the lemma, p. 414, we obtain (36).

STANFORD UNIVERSITY.

REFERENCES.

[A. 1] A. Ahlfors, "Untersuchungen zur Theorie der konformen Abbildung und der ganzen Funktionen," *Acta Societatis scientiarum Fennicae*, N. s. 1, Nr. 9, 1930.

[A. 2] ———, "On quasiconformal mappings," *Journal d'Analyse Mathématique*, vol. 3 (1954), pp. 1-58 and 207-208.

[B. 1] P. P. Belinskij, "On distortion and quasi-conformal mappings," *Doklady Akademii Nauk SSSR* 91 (1933).

[B. 2] S. Bergman, "Some bounds for coefficients of schlicht pseduo-conformal mappings," *Annali di Matematica Pura ed Applicata*, (IV), 57 (1962), pp. 295-310.

[B. 3] ———, "Sur quelques propriétés des transformations par un couple des fonctions de deux variables complexes," *Atti della Accademia Nazionali dei Lincei, Rendiconti*, 6e ser., 19 (1934), pp. 474-478.

[B. 4] ———, "A formula for the stream function of certain flows," *Proceedings of the National Academy of Sciences, U. S. A.*, 29 (1943), pp. 276-281.

[B. 5] ———, "Functions in the extended class in the theory of the theory of functions of several complex variables," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 423-447.

[B. 6] ———, *The kernel function and conformal mapping*. Mathematical Surveys 5, American Mathematical Society, 1950.

[B. 7] ———, "Kernel function and extended classes in the theory of functions of complex variables," *Colloque sur les fonctions de plusieurs variables*, Centre de Recherches Math., Brussels (1953), pp. 135-157.

[B. 8] ———, "On representation of stream functions of subsonic and supersonic flows of compressible fluids," *Journal of Rational Mechanics and Analysis*, vol. 4 (1955), pp. 883-905.

[B. 9] ———, *Sur les fonctions orthogonales de plusieur variables complexes avec les applications à la théorie des fonctions analytiques*, Interscience Publishers, New York, 1941, and Mémorial des Sciences Mathématiques, 102, Paris, 1947.

[B. 10] ———, "A class of pseudo-conformal and quasi-conformal mappings," *Mathematische Annalen*, vol. 136 (1958), pp. 134-138.

[B. 11] ———, "A class of quasi-conformal transformations in the theory of functions of two complex variables," *Journal of Mathematics and Mechanics*, vol. 7 (1958), pp. 937-956.

[B. 12] L. Bers, "On the continuation of a potential gas flow across the sonic line," NACA, Tech. Note 2058, 1950.

[B. 13] ———, *Theory of pseudonanalytic functions*. Lecture Notes, New York University, New York, 1953.

[B. 14] ———, *Function-theoretical properties of solutions of partial differential equations of elliptic type*. Annals of Mathematics Studies 33 (1954), pp. 69-94.

[B. 15] ———, "An outline of the theory of pseudoanalytic functions," *Bulletin of the American Mathematical Society*, vol. 62 (1956), pp. 291-331.

[B. 16] ———, "Mathematical aspects of subsonic and transonic gas dynamics," *Surveys in Applied Mathematics*, III, John Wiley and Sons, New York, 1958.

[B. 17] H. Bremermann, "Holomorphic continuation of the kernel function and the Bergman metric in several complex variables," *Lectures on functions of a complex variable*, The University of Michigan Press (1955), pp. 349-383.

[B. G. 1] L. Bers and A. Gelbart, "On a class of differential equations in mechanics of continua," *Quarterly of Applied Mathematics*, vol. 1 (1943), pp. 168-188.

[B. G. 2] ———, "On a class of functions defined by partial differential equations," *Transactions of the American Mathematical Society*, vol. 56 (1944), pp. 67-93.

[H. 1] S. Hitotumatu, "Some remarks on a quasi-pseudo-conformal mapping of Reinhardt circular domains," *Commentariorum Mathematicorum Universitatis Sancti Pauli*, Tokyo, vol. 8 (1960), pp. 1-21.

[L. 1] C. Loewner, "Some bounds for the critical free stream Mach number of a compressible flow around an obstacle," *Studies in Mathematics and Mechanics*, Academic Press Inc., New York, 1954, pp. 177-183.

[L. 2] ———, "Conservation laws in compressible fluid flow and associated mappings," *Journal of Rational Mechanics and Analysis*, vol. 2 (1953), pp. 537-561.

[L. 3] D. B. Lowdenslager, "Potential theory in bounded symmetric homogeneous complex domains," *Annals of Mathematics*, vol. 67 (1958), pp. 467-484.

[L. 4] ———, "Potential theory and a generalized Jensen-Nevanlinna formula for functions of several complex variables," *Journal of Mathematics and Mechanics*, vol. 7 (1958), pp. 207-218.

[N. 1] R. Nevanlinna, *Eindeutige analytische Funktionen*, Grundlehren der Mathematische Wissenschaften, 46, Springer, Berlin, 1936.

NORMAL CONGRUENCE SUBGROUPS OF THE MODULAR GROUP.*[1]

By Morris Newman.

1. Introduction. Let Γ be the 2×2 modular group; that is, the group of 2×2 rational integral matrices of determinant 1 in which a matrix is identified with its negative. If n is a positive integer, $\Gamma(n)$ denotes the *principal congruence subgroup* of Γ of *level* n, consisting of all elements of Γ congruent modulo n to a scalar matrix. The subgroup of $\Gamma(n)$ consisting of all elements of Γ congruent modulo n to the identity matrix I is denoted by $\Gamma_1(n)$. Then $\Gamma(n)$, $\Gamma_1(n)$ are normal subgroups of Γ. A subgroup G of Γ containing a group $\Gamma(n)$ is termed a *congruence subgroup*, and is said to be of *level* n if n is the least such integer. Notice that $\Gamma_1(n)$ is not in general a congruence subgroup, according to the definition above. The groups $\Gamma_1(n)$ are fully discussed in Gunning's book [2] and it should be noted that these are customarily referred to as principal congruence subgroups.

The principal result of this paper is that *a normal congruence subgroup of level n, where $(n, 6) = 1$ is necessarily the principal congruence subgroup* $\Gamma(n)$. The proof will be arranged for an induction and will in fact yield somewhat more. At the end of the paper some open questions are discussed.

2. Preliminary material. If m and n are positive integers, then (m, n) will stand for their greatest common divisor and $[m, n]$ for their least common multiple. We set

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Then, as is well-known, S and T generate Γ and

$$T^2 = (ST)^3 = -I$$

are defining relations for Γ.

We require the following lemma:

Lemma 1. *The principal congruence subgroups satisfy*

$$\Gamma(m)\Gamma(n) = \Gamma((m, n)), \tag{1}$$

$$\Gamma(m) \cap \Gamma(n) = \Gamma([m, n]). \tag{2}$$

* Received ————————; revised July 9, 1963.

[1] The preparation of this paper was supported by the Office of Naval Research.

Hence by one of the isomorphism theorems

$$\Gamma((m,n))/\Gamma(m) \cong \Gamma(n)/\Gamma([m,n]). \tag{3}$$

Stated otherwise, (3) becomes

$$\Gamma(d)/\Gamma(da) \cong \Gamma(db)/\Gamma(dab), \tag{4}$$

where d is arbitrary and $(a,b)=1$. A proof of Lemma 1 for the groups $\Gamma_1(n)$ is given in [4]. The modifications required for the groups $\Gamma(n)$ are slight, and we forego the proof.

The index $(\Gamma\colon\Gamma(n))$ is readily computed: It is just

$$(\Gamma\colon\Gamma_1(n))/(\Gamma(n):\Gamma_1(n))$$

where

$$(\Gamma\colon\Gamma_1(n)) = n^3\prod_{p\mid n}(1-1/p^2), \tag{5}$$

$$(\Gamma(n):\Gamma_1(n)) = 2^{w(n)}f(n). \tag{6}$$

Here $w(n)$ is the number of distinct primes dividing n, and

$$f(n) = \begin{cases} 1 & n \text{ odd} \\ \frac{1}{2} & 2 \parallel n \\ 1 & 4 \parallel n \\ 2 & 8 \mid n. \end{cases} \tag{7}$$

We note that if p is an odd prime dividing n, then these imply

$$(\Gamma(n):\Gamma(np)) = p^3. \tag{8}$$

When n is itself a power of an odd prime, $\Gamma(n)$ consists just of those matrices $M\in\Gamma$ satisfying $M\equiv\pm I \pmod{n}$. We now prove the following lemmas:

LEMMA 2. *Let p be an odd prime, $n>1$ a power of p. Then $\Gamma(n)/\Gamma(np)$ is an abelian group of order p^3 and of type (p,p,p). The generators may be chosen modulo $\Gamma(np)$ as*

$$A=\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix},\quad B=\begin{pmatrix}1 & 0\\ n & 1\end{pmatrix},\quad C=\begin{pmatrix}1+n & -n\\ n & 1-n\end{pmatrix}. \tag{9}$$

LEMMA 3. *Let p be an odd prime, $p\mid n$. Then $\Gamma(n)/\Gamma(np)$ is an abelian group of order p^3 and of type (p,p,p).*

Proof of Lemmas 2 and 3. We first prove that Lemma 3 is a direct consequence of Lemma 2. We can write $n=p^t n_1$, where $t\geqq 1$ and $(n_1,p)=1$.

Then

$$\Gamma(n)/\Gamma(np) = \Gamma(n_1p^t)/\Gamma(n_1p^{t+1})$$
$$\cong \Gamma(p^t)/\Gamma(p^{t+1}),$$

by (4). By Lemma 2 the latter is abelian of order p^3 and of type (p, p, p) and Lemma 3 follows.

We turn now to the proof of Lemma 2. Suppose that $U, V \in \Gamma(n)$. Then $U \equiv uI \pmod{n}$, $V \equiv vI \pmod{n}$, where $u, v = \pm 1$. Thus

$$(U - uI)(V - vI) \equiv 0 \pmod{n^2}$$

and it follows that

$$UV \equiv vU + uV - uvI \equiv VU \pmod{n^2}.$$

Since $p \mid n$ the congruence also holds modulo np which implies that $\Gamma(n)/\Gamma(np)$ is abelian. Now let

$$M = \pm \begin{pmatrix} 1 + \alpha n & \beta n \\ \gamma n & 1 + \delta n \end{pmatrix}$$

be any element of $\Gamma(n)$. Then $\alpha + \delta \equiv 0 \pmod{n}$, and it is easily verified that

$$M \equiv \pm A^{\alpha+\beta} B^{\gamma-\alpha} C^{\alpha} \pmod{n^2}.$$

Hence the matrices (9) are certainly generators modulo $\Gamma(np)$, and each is of period p modulo $\Gamma(np)$. That they are independent modulo $\Gamma(np)$ is seen from the congruence

$$A^x B^y C^z \equiv \begin{pmatrix} 1 + nz & n(x - z) \\ n(y + z) & 1 - nz \end{pmatrix} \pmod{n^2}$$

which implies that $A^x B^y C^z \equiv \pm I \pmod{np}$ if and only if $x \equiv y \equiv z \equiv 0 \pmod{p}$. The proof of the lemma is complete. We remark that when $n = p = 2$ the matrices A, B, C are not independent, since then $ABC \equiv -I \pmod{4}$.

3. The principal lemmas. In this section we develop the lemmas directly required in the proof of the main theorem. It will be necessary to give special proofs for a prime or the square of a prime to make the induction go through generally, and the lemmas that follow are devoted to this object. Throughout this section m, n denote positive integers and p a prime.

LEMMA 4. *Let p be a prime > 3. Let G be a normal subgroup of Γ such that*

$$\Gamma \supset G \supset \Gamma(p).$$

Then $G = \Gamma$ or $\Gamma(p)$.

Proof. This lemma is proved by Dickson in his book [1] on the linear groups, in another form. Dickson shows that $LF(2,p) \cong \Gamma/\Gamma(p)$ is a simple group, if p is a prime > 3.

We introduce the notation (to be used in the lemma that follows)

$$A^x B^y C^z = (x, y, z) \tag{10}$$

where A, B, C are given by (9).

LEMMA 5. *Let p be a prime > 2, $p \mid n$. Let G be a normal subgroup of Γ such that*

$$\Gamma(n) \supset G \supset \Gamma(np).$$

Then $G = \Gamma(n)$ or $\Gamma(np)$.

Proof. Assume first that n is a power of p. We remark that $G/\Gamma(np)$ is abelian since it is a subgroup of the abelian group $\Gamma(n)/\Gamma(np)$ (Lemma 2). Suppose that $G \neq \Gamma(np)$. Then there is an element $M = (x, y, z) \in G$ where at least one of x, y, z is prime to p. Since G is normal,

$$S(M) = SMS^{-1}, \qquad T(M) = TMT^{-1}$$

belong to G. It is an easy computation that

$$S(A) = A, \quad S(B) = C, \quad S(C) \equiv A^{-2}B^{-1}C^{2} \pmod{n^2} \tag{11}$$

and that

$$T(A) = B^{-1}, \quad T(B) = A^{-1}, \quad T(C) \equiv A^{-2}B^{2}C^{-1} \pmod{n^2}. \tag{12}$$

We find from (11) therefore that

$$\begin{gathered} M = (x, y, z) \in G, \\ M_1 = (x - 2z, -z, y + 2z) \in G, \\ M_2 = (x - 2y - 6z, -y - 2z, 2y + 3z) \in G. \end{gathered}$$

Thus $MM_1^{-2}M_2 = (-2y - 2z, 0, 0) \in G$ which implies that $(y + z, 0, 0) \in G$ since $p > 2$. There are now two possibilities: If $(y + z, p) = 1$ then $(1, 0, 0) = A \in G$. The equations (11), (12) now imply that $(0, 1, 0) = B$ and $(0, 1, 1) = C$ also belong to G which in turn implies that $G = \Gamma(n)$. If however $y + z \equiv 0 \pmod{p}$ then from $M^{-1}M_1$ we deduce that $(2y, 0, 0) \in G$ which implies that $(y, 0, 0) \in G$ since $p > 2$. If $p \mid y$ then also $p \mid z$ and so $(x, p) = 1$, $(x, 0, 0) \in G$ which implies that $(1, 0, 0) \in G$. As before this gives $G = \Gamma(n)$. If $(p, y) = 1$ then also $(1, 0, 0) \in G$ again implying $G = \Gamma(n)$. Thus in all cases we have proved that $G = \Gamma(n)$ or $\Gamma(np)$, and the proof of the lemma is concluded, if n is a power of p.

Now suppose that $n = p^t n_1$ where $t \geqq 1$ and $(n_1, p) = 1$. Then the inclusions of the lemma become

(13) $$\Gamma(p^t n_1) \supset G \supset \Gamma(p^{t+1} n_1).$$

Intersecting and producting in (13) with $\Gamma(p^{t+1})$, we find that

$$\Gamma(p^{t+1} n_1) \supset G \cap \Gamma(p^{t+1}) \supset \Gamma(p^{t+1} n_1),$$
$$\Gamma(p^t) \supset G\Gamma(p^{t+1}) \supset \Gamma(p^{t+1}).$$

Hence $G \cap \Gamma(p^{t+1}) = \Gamma(p^{t+1} n_1)$, and (by the first part of the lemma) $G\Gamma(p^{t+1}) = \Gamma(p^t)$ or $\Gamma(p^{t+1})$. If $G\Gamma(p^{t+1}) = \Gamma(p^{t+1})$ then $G \subset \Gamma(p^{t+1})$. Thus

$$G \subset \Gamma(p^{t+1}) \cap \Gamma(p^t n_1) = \Gamma(p^{t+1} n_1)$$
$$\Gamma(p^t) \supset G\Gamma(p^{t+1}) \supset \Gamma(p^{t+1}.$$

which implies that $G = \Gamma(p^{t+1} n_1)$. Thus we can assume that $G\Gamma(p^{t+1}) = \Gamma(p^t)$. Then (3) gives

$$\Gamma(p^t)/\Gamma(p^{t+1}) \cong G/\Gamma(p^{t+1} n_1).$$

But also by (3)

$$\Gamma(p^t)/\Gamma(p^{t+1}) \cong \Gamma(p^t n_1)/\Gamma(p^{t+1} n_1).$$

Hence

$$\Gamma(p^t n_1)/\Gamma(p^{t+1} n_1) \cong G/\Gamma(p^{t+1} n_1)$$

and since $G \subset \Gamma(p^t n_1)$, it follows that $G = \Gamma(p^t n_1)$. This completes the proof of Lemma 5.

We also require

LEMMA 6. *Let p be a prime > 3. Let G be a normal subgroup of Γ such that*

(14) $$\Gamma \supset G \supset \Gamma(p^2).$$

Then $G = \Gamma$, $\Gamma(p)$ or $\Gamma(p^2)$.

Proof. Intersecting and producting in (14) with $\Gamma(p)$, we find that

$$\Gamma(p) \supset G \cap \Gamma(p) \supset \Gamma(p^2),$$
$$\Gamma \supset G\Gamma(p) \supset \Gamma(p).$$

Lemma 5 implies that $G \cap \Gamma(p) = \Gamma(p)$ or $\Gamma(p^2)$, and Lemma 4 implies that $G\Gamma(p) = \Gamma$ or $\Gamma(p)$. If $G \cap \Gamma(p) = \Gamma(p)$ then $\Gamma \supset G \supset \Gamma(p)$, $G = \Gamma$ or $\Gamma(p)$. If $G\Gamma(p) = \Gamma(p)$ then $\Gamma(p) \supset G \supset \Gamma(p^2)$, $G = \Gamma(p)$ or $\Gamma(p^2)$. Assume then that $G \cap \Gamma(p) = \Gamma(p^2)$, $G\Gamma(p) = \Gamma$. Then $\Gamma/G \cong \Gamma(p)/\Gamma(p^2)$. By Lemma 2 $\Gamma(p)/\Gamma(p^2)$ is abelian and so therefore is Γ/G. This implies that

$G \supset \Gamma'$, the commutator subgroup of Γ. But $(\Gamma:\Gamma') = 6$ (see [3]) and so $(\Gamma: G) \mid 6$. However

$$(\Gamma: G) = (\Gamma(p) : \Gamma(p^2)) = p^3,$$

and so $p^3 \mid 6$. This is impossible however. We have proved therefore that always, $G = \Gamma$, $\Gamma(p)$ or $\Gamma(p^2)$ and the proof of the lemma is completed.

We also prove

LEMMA 7. *Let p be a prime > 2, $p \mid n$. Let G be a normal subgroup of Γ such that*

$$\Gamma(n) \supset G \supset \Gamma(np^2). \tag{15}$$

Then $G = \Gamma(n)$, $\Gamma(np)$ or $\Gamma(np^2)$.

Proof. Intersecting and producting in (15) with $\Gamma(np)$, we find that

$$\Gamma(np) \supset G \cap \Gamma(np) \supset \Gamma(np^2),$$
$$\Gamma(n) \supset G\Gamma(np) \supset \Gamma(np).$$

Then Lemma 5 implies that $G \cap \Gamma(np) = \Gamma(np)$ or $\Gamma(np^2)$, and that $G\Gamma(np) = \Gamma(n)$ or $\Gamma(np)$. If $G \cap \Gamma(np) = \Gamma(np)$ then $\Gamma(n) \supset G \supset \Gamma(np)$ and $G = \Gamma(n)$ or $\Gamma(np)$. If $G\Gamma(np) = \Gamma(np)$ then $\Gamma(np) \supset G \supset \Gamma(np^2)$ and $G = \Gamma(np)$ or $\Gamma(np^2)$. Assume then that $G \cap \Gamma(np) = \Gamma(np^2)$, $G\Gamma(np) = \Gamma(n)$. Then

$$G/\Gamma(np^2) \cong \Gamma(n)/\Gamma(np).$$

Since $p \mid n$ $\Gamma(n)/\Gamma(np)$ is abelian of type (p,p,p) (Lemma 3) and so therefore is $G/\Gamma(np^2)$. Suppose that P, Q, R are elements of G (and hence of $\Gamma(n)$) such that $G/\Gamma(np^2)$ is generated by P, Q, R modulo $\Gamma(np^2)$. Then we have $P = \mu I + nE$, E an integral matrix. However, we know that $P^p \in \Gamma(np^2)$ and

$$P^p = (\mu I + nE)^p \equiv \mu^p I + \mu^{p-1} npE \pmod{np^2},$$

since $p \mid n$. This implies that $P \in \Gamma(np)$. Similarly Q and R belong to $\Gamma(np)$ and so $G \subset \Gamma(np)$. Hence $G = \Gamma(np)$, and the lemma follows.

LEMMA 8. *Suppose m has the property that for any normal subgroup G of Γ*

$$\Gamma \supset G \supset \Gamma(m) \text{ implies } G = \Gamma(d), d \mid m.$$

Then m also has the property that for any normal subgroup G of Γ and any n with $(m,n) = 1$

$$\Gamma(n) \supset G \supset \Gamma(mn) \text{ implies } G = \Gamma(dn), d \mid m.$$

Proof. Let G be a normal subgroup of Γ such that

$$\Gamma(n) \supset G \supset \Gamma(mn).$$

Producting by $\Gamma(m)$ we obtain

$$\Gamma \supset G\Gamma(m) \supset \Gamma(m).$$

By the hypotheses of the lemma it follows that

$$G\Gamma(m) = \Gamma(d), d \mid m.$$

Now intersecting by $\Gamma(m)$ we obtain

$$\Gamma(mn) \supset G \cap \Gamma(m) \supset \Gamma(mn),$$
$$G \cap \Gamma(m) = \Gamma(mn).$$

Hence

$$\Gamma(d)/\Gamma(m) \cong G/\Gamma(mn).$$

But also by (4)

$$\Gamma(d)/\Gamma(m) \cong \Gamma(dn)/\Gamma(mn),$$

since $(m, n) = 1$ and $d \mid m$. It follows that

$$G/\Gamma(mn) \cong \Gamma(dn)/\Gamma(mn). \tag{16}$$

But $\Gamma(d) \supset G$ (since $\Gamma(d) = G\Gamma(m)$) and $\Gamma(n) \supset G$. It follows that $\Gamma(dn) \supset G$ (since $(m, n) = 1$ and $d \mid m$) and this together with (16) implies that $\Gamma(dn) = G$, completing the proof of the lemma.

4. The principal theorem. We combine Lemmas 4, 5, 6, 7, 8 in our first theorem:

THEOREM 1. *Let p be a prime > 3, n a positive integer. If G is a normal subgroup of Γ such that*

$$\Gamma(n) \supset G \supset \Gamma(np)$$

then $G = \Gamma(n)$ or $\Gamma(np)$; and if G is a normal subgroup of Γ such that

$$\Gamma(n) \supset G \supset \Gamma(np^2)$$

then $G = \Gamma(n)$, $\Gamma(np)$ or $\Gamma(np^2)$.

We are now in a position to prove

THEOREM 2 (*the principal theorem*). *Let m, n be positive integers, $(m, 6) = 1$. Let G be a normal subgroup of Γ such that*

$$\Gamma(n) \supset G \supset \Gamma(mn). \tag{17}$$

Then $G = \Gamma(nd)$, $d \mid m$.

Proof. The proof will be by induction on n and on $\Omega(m)$, the total number of primes dividing m.

The theorem is certainly true for all n and for $m=1$, p or p^2 where p is a prime >3, this being the content of Theorem 1. Suppose the theorem true for all n and for all m such that $(m,6)=1$, $\Omega(m)<k$. Let m be such that $(m,6)=1$, $\Omega(m)=k$. We can assume that m is neither a prime nor the square of a prime since the theorem has already been proved in these cases. Let d be any divisor of m, $d>1$. Intersecting by $\Gamma(nd)$ in (17) we obtain

$$\Gamma(nd)\supset G\cap\Gamma(nd)\supset\Gamma(mn).$$

By the induction hypothesis (with nd replacing n)

$$G\cap\Gamma(nd)=\Gamma(nd\bar{d}),\bar{d}\mid m/d.$$

There are two possibilities. If $d\bar{d}<m$ for some d, then $\Omega(d\bar{d})<\Omega(m)$, $\Gamma(n)\supset G\supset\Gamma(nd\bar{d})$ and the proof is completed by induction. Otherwise we have that

$$G\cap\Gamma(nd)=\Gamma(mn),d\mid m,d>1. \tag{18}$$

We now apply the same process with products. Producting by $\Gamma(nd)$ in (17) we obtain

$$\Gamma(n)\supset G\Gamma(nd)\supset\Gamma(nd).$$

Then if $d<m$, the induction hypothesis implies that

$$G\Gamma(nd)=\Gamma(n\delta),\delta\mid d.$$

Thus

$$\Gamma(n\delta)\supset G\supset\Gamma(mn).$$

Again there are two possibilities. If $\delta>1$ for some d, then the proof is completed by induction. Otherwise we have that

$$G\Gamma(nd)=\Gamma(n),d\mid m,d<m. \tag{19}$$

Combining (18) and (19) we conclude that for every divisor d of m, $1<d<m$ we have

$$G/\Gamma(mn)\cong\Gamma(n)/\Gamma(nd).$$

Hence $(\Gamma(n):\Gamma(nd))$ is *independent of the choice* of d, $1<d<m$. But m is larger than 1 and is neither a prime nor the square of a prime. This implies that primes p, q (not necessarily distinct) exist such that $pq\mid m$, $pq<m$. Then choosing $d=p$, $d=pq$ we find that

$$(\Gamma(n):\Gamma(np))=(\Gamma(n):\Gamma(npq))$$

so that

$$(\Gamma(np) : \Gamma(npq)) = 1.$$

This is not possible however (since $\Gamma(np) \neq \Gamma(npq)$) and so the proof of the theorem is concluded.

5. Some open questions. The restriction $(m, 6) = 1$ cannot in general be removed. For example if Γ^n is the fully invariant subgroup of Γ generated by the n-th powers of the elements of Γ then

$$\Gamma \supset \Gamma^2 \supset \Gamma(2),$$
$$\Gamma \supset \Gamma^3 \supset \Gamma(3),$$
$$\Gamma \supset \Gamma' \supset \Gamma(6).$$

Here Γ^2, Γ^3 and Γ' (the commutator subgroup of Γ and just $\Gamma^2 \cap \Gamma^3$) are normal congruence subgroups of levels 2, 3, 6 respectively but are not principal congruence subgroups. Similarly the intersections $\Gamma^2 \cap \Gamma(n)$, $\Gamma^3 \cap \Gamma(n)$, $\Gamma' \cap \Gamma(n)$ are normal congruence subgroups but not principal congruence subgroups. The problem of classifying all normal congruence subgroups of level n where $(n, 6) > 1$ remains open. The theorem can be extended to the higher dimensional modular groups, but the extension to the symplectic modular group is (presumably) more difficult and is also an open question.

REFERENCES.

[1] L. E. Dickson, *Linear groups*, New York, 1958.

[2] R. C. Gunning, *Lectures on modular forms*, Princeton, 1962.

[3] M. Newman, "The structure of some subgroups of the modular group," *Illinois Journal of Mathematics*, vol. 6 (1962), pp. 480-487.

[4] M. Newman and J. R. Smart, "Modulary groups of $t \times t$ matrices," *Duke Mathematical Journal*, vol. 30 (1963), pp. 253-257.

ON AFFINE TRANSFORMATIONS OF COMPACT ABELIAN GROUPS.*

By F. J. Hahn.[1]

Introduction. By an affine transformation of a group we mean a homeomorphism of the group onto itself which is the composition of an automorphism followed by a translation. The purpose of this paper is to study the ergodic theory of affine transformations of compact abelian groups. In Section 1 such transformations are completely classified according to whether they are ergodic, weakly mixing or strongly mixing. In Section 2, we concentrate on the affine transformations of the torus and classify these according to the eigenvalues of the automorphism part of the affine transformation.[2] Section 3 makes a closer examination of the spectrum of an affine transformation. The L_2 space of the group is decomposed into two stable orthogonal subspaces on one of which the affine transformation has discrete spectrum and on the other of which the affine transformation has continuous spectrum. In Section 4 the question of strict ergodicity is discussed and standard forms are derived for transformation of the torus. It is shown that a certain class of ergodic affine transformations of the torus are always strictly ergodic. In Section 5 we apply some of the previous results (Section 2 and 4) to prove the following theorem of Hermann Weyl [8]: Let

$$p(z) = a_m z^m + a_{m-1} z^{m-1} + \cdots + a_1 z + a_0$$

be a polynomial with real coefficients and $a_m \neq 0$, $m > 0$. If there is an i, $m \geqq i > 0$, for which a_i is irrational then the numbers $p(n)$, $n = 0, 1, 2, 3 \cdots$ are uniformly distributed mod 1.

1. Affine transformations of compact Abelian groups. If X is a compact abelian group we let $\hat{X}$ be its character group. The elements of X shall be designated by lower case Roman letters $x, y, z \cdots$ and the elements of $\hat{X}$ shall be designated by lower case Greek letters $\xi, \eta, \zeta, \cdots$. We use the additive notation for X and indicate the value of the character ξ at x by $\xi(x)$. By $\mathrm{Aut}(X)$ we mean the automorphism group of X. The group $\mathrm{Aff}(X)$

* Received August 1, 1963.

[1] The author was partially supported by NSF G 25222.

[2] The results of this section have been announced in L. Auslander and F. Hahn, "Discrete transformations on tori and flows on Solvmanifolds," *Bulletin of the American Mathematical Society*, vol. 68 (1962), pp. 614-615.

$= X \cdot \mathrm{Aut}(X)$ will be called the *group of affine transformations of* X. Here the dot indicates semi direct product. An element (x_0, A) in $\mathrm{Aff}(X)$ acts on X as follows $(x_0, A)x = x_0 + A(x)$. Since $\mathrm{Aut}(X)$ preserves Haar measure we see that each element $S \in \mathrm{Aff}(X)$ is measure preserving. Each such S induces a unitary operator V_S on $L_2(X)$ as follows $(V_S f)(x) = f(S(x))$. We recall that the character group X forms an ortho-normal basis for $L_2(X)$.

We list the following definitions and results from ergodic theory:

1. The transformation S is said to be ergodic if the only eigenfunctions with eigenvalue one of V_S are the constant functions.

2. The transformation S is said to be weakly mixing if the only eigenfunctions of V_S are the constant functions.

3. The transformation S is said to be strongly mixing if for each $f, g \in L_2(X)$ we have $\lim_n (V_S{}^n f, g) = (f, 1)(1, g)$.

We recall that in general strongly mixing implies weakly mixing implies ergodic. (For instance see [6] or [7]) If $S = (x_0, A)$ and $x_0 = 0$ then Halmos [5] has shown that ergodic, weakly mixing, and strongly mixing are equivalent. He has shown that A is ergodic precisely when the transformation V_A which it induces on $\hat{X}$ has no periodic orbits aside from the trivial one. On the opposite hand when A is the identity and $x_0 \in X$ we know that S is never weakly mixing. We also know that S is ergodic if and only if x_0 generates a dense subgroup of X. In this section we shall show that for $S \in \mathrm{Aff}(X)$ ergodic is not equivalent to weakly mixing but that weakly mixing and strongly mixing are equivalent. We shall also characterize these cases completely.

THEOREM 1. *If* $S = (x_0, A) \in \mathrm{Aff}(X)$ *and if each character* $\xi \neq 1$ *has an infinite orbit under* V_A *then* S *is strongly mixing.*

Proof. If $\xi = 1$ then

$$1 = (\xi, 1)(1, \xi) = (\xi, \xi) = \lim_n (V_S{}^n \xi, \xi).$$

If $\xi \neq 1$ then we observe that $V_S(\xi) = \xi(x_0) V_A \xi$ and in general $V_S{}^n(\xi) = \lambda_n(x_0) V_A{}^n \xi$ where $|\lambda_n(x_0)| = 1$. Since $V_A{}^n \xi$ is again a character it follows from the hypothesis that it can never be a particular character for more than one value of n. We have

$$(\xi, 1)(1, \eta) = 0 = \lim_n \lambda_n(\xi)(V_A{}^n \xi, \eta) = \lim_n (V_S{}^n \xi, \eta).$$

We have shown that if H_0 is the subspace of $L_2(X)$ consisting of constant functions then V^n converges weakly to the identity on H_0. On the other hand $(V_S{}^n\xi, \eta)$ converges to zero if $\xi \not\equiv 1$ and $\xi, \eta \in \hat{X}$. Using the fact that any elements f, g in $L_2(X)$ have Fourier series expansions in terms of the characters and standard continuity and linearity arguments we obtain $(f, 1)(1, g) = \lim_n (V_S{}^n f, g)$ showing that S is strongly mixing.

THEOREM 2. *If $S = (x_0, A) \in \mathrm{Aff}(X)$ and if there is a $\xi \in \hat{X}$, $\xi \not\equiv 1$ and a positive integer n for which $V_A{}^n\xi = \xi$ then S is not weakly mixing. (Consequently S is not strongly mixing.)*

Proof. Choose $\xi \in \hat{X}$, $\xi \not\equiv 1$ and n satisfying the above hypotheses. We see that $V_S{}^n\xi = \lambda(\xi)\xi$ where $|\lambda(\xi)| = 1$. Let M be the space spanned by $\xi, V_S\xi, \cdots, V_S{}^{n-1}\xi$. We see that M is orthogonal to the constant functions and $V_S(M) \subset M$. Since M is finite dimensional it follows that there is an eigenfunction $g \in M$ for V_S. Since g is not constant we see that S is not weakly mixing.

COROLLARY 3. *If $S = (x_0, A) \in \mathrm{Aff}(X)$ then the following statements are equivalent.*

1. The unitary operator V_A restricted to $\hat{X}$ has no periodic points except the trivial one.

2. A is ergodic.

3. S is strongly mixing.

4. S is weakly mixing.

The equivalence of 1, 3, 4 is the content of the preceding theorems. The equivalence of 1 and 2 is proved in [5] by Halmos.

We now wish to examine the case in which A is not ergodic. If A is not ergodic then then set F of characters ξ, $\xi \not\equiv 1$, which are periodic under V_A is not empty. For any non-negative integer n we see $V_S{}^n\xi = \lambda_n(\xi) V_A{}^n\xi$ where $\lambda_n(\xi) = \xi(x_0 + Ax_0 + A^2x_0 + \cdots + A^{n-1}x_0)$.

LEMMA 4. *If $S = (x_0, A) \in \mathrm{Aff}(X)$ and if there is a character $\xi, \xi \not\equiv 1$ and a positive integer n such that $V_S{}^n\xi = \xi$ then S is not ergodic.*

Proof. Let $g = \xi + V_S\xi + V_S{}^2\xi + \cdots + V_S{}^{n-1}\xi$. Since $(g, 1) = 0$ we see that g is not constant. Since $V_S g = g$ we see that S is not ergodic.

THEOREM 4. *Suppose $S = (x_0, A) \in \mathrm{Aff}(X)$ and S is not strongly mixing.*

Under these conditions S is ergodic if and only if for each $\xi \in F$ and for each positive integer n for which $V_A{}^n\xi = \xi$ we have $\lambda_n(\xi) \neq 1$.

Proof. If $\xi \in F$ and if $V_A{}^n\xi = \xi$ and $\lambda_n(\xi) = 1$ we have $V_S{}^n\xi = \lambda_n(\xi) V_A{}^n(\xi) = \xi$ and it follows from Lemma 4 that S is not ergodic.

Suppose now that for each $\xi \in F$ and for each n for which $V_A{}^n\xi = \xi$ we have $\lambda_n(\xi) \neq 1$. Under these conditions we wish to show that if $g \in L_2(X)$ and $V_S g = g$ then g is a constant. Since the elements of $\hat{X}$ form an orthonormal basis for $L_2(X)$ we expand g in a Fourier series relative to this basis and we obtain

$$g = \sum_i (g, \xi_i) \xi_i.$$

Also

$$V_S g = \sum_i (g, \xi_i) V_S \xi_i = \sum_i (g, \xi_i) \lambda_1(\xi_i) V_A \xi_i,$$

For each $\xi_i \in \hat{X}$ we have $V_A \xi_i \in \hat{X}$. Comparing coefficients of like terms in g and $V_S g$ we obtain

$$(g, V_A \xi_i) = \lambda_1(\xi_i)(g, \xi_i)$$

and consequently

$$|(g, V_A \xi_i)| = |(g, \xi_i)|.$$

We wish to show that this implies $(g, \xi_i) = 0$ if $\xi_i \not\equiv 1$. We distinguish two cases $\xi_i \notin F$, $\xi_i \not\equiv 1$; and $\xi_i \in F$. In the first case if $(g, \xi_i) \neq 0$ we would obtain $0 \neq |(g, \xi_i)| = |(g, V_A{}^n \xi_i)|$ for each n. This is impossible since $\sum_i |(g, \xi_i)|^2 < +\infty$.

Suppose now $\eta \in F$. There is an n for which $V_A{}^n \eta = \eta$. We also have $g = V_S{}^n g = \sum_i (g, \xi_i) V_S{}^n \xi_i = \sum_i (g, \xi_i) \lambda_n(\xi_i) V_A{}^n \xi_i$. The character η appears somewhere in this sum and $V_A{}^n \eta = \eta$. Comparing coefficients of g and $V_S{}^n g$ we obtain

$$(g, \eta) = \lambda_n(\eta)(g, \eta).$$

Since $\lambda_n(\eta) \neq 1$ we must have $(g, \eta) = 0$.

Since all other coefficients vanish we have $g = (g, 1)$ and thus S is ergodic.

We remark that in the special case where A is the identity the above theorem yields the classical result which states that S is ergodic if and only if $\xi(x_0) = 1$ implies $\xi \equiv 1$ for each $\xi \in \hat{X}$.

2. Affine transformations of the torus. We wish to apply the pre-

vious results to classify affine transformations of the torus. We let X be the m dimensional torus. Then $X = R^m/D$ where R^m is the additive group of the real Euclidean m dimensional vector space and D is the subgroup of points $n = (n_1, \cdots, n)$ where n_i are integers. Any automorphism A of X may be lifted in a unique fashion to a linear transformation of R^m which preserves D. Conversely any non-singular linear transformation A of R^m which preserves D induces a unique automorphism of X. If we let $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ (1 in the i-th place) $i = 1, 2, \cdots, m$, be a basis for R^m then any non-singular linear transformation A of R^m which preserves D has a matrix representation (a_{ij}), with respect to the e_i, in which a_{ij} are integers and determinant $(a_{ij}) = \pm 1$. We shall denote the adjoint of A by A^*.

The character group $\hat{X}$ is isomorphic to D and this isomorphism is given by $(n_1, n_2, \cdots, n_m) = n \to \exp(x, n)$ where $\exp\theta = e^{2\pi i\theta}$ and $(x, n) = \sum_{i=1}^{m} n_i x_i$ where $x = (x_1, x_2, \cdots, x_m)$. We note that if $A \in \operatorname{Aut}(X)$ then $V_A \exp(x, n) = \exp(Ax, n) = \exp(x, A^*n)$. Here A is looked upon as a linear transformation of R^m which preserves D.

For each positive integer p we define

$$K_p(A) = \operatorname{Kernel}[A^{*p} - I] \cap D.$$

Since $A^* - I$ is a factor of $A^{*p} - I$ we see that $K_1(A) \subset K_p(A)$ for $p \geqq 1$. We now interpret the theorems of the previous section in terms of the eigenvalues of A or what is the same in terms of $K_p(A)$.

THEOREM 5. *If $S = (x_0, A)$ is an affine transformation of the m dimensional torus X and if $K_p(A) = \{0\}$ for all $p \geqq 1$ then S is strongly mixing.*

Proof. Because of Theorem 1 we need only show that each non-trivial character has an infinite orbit under V_A. If this were not so there would exist an $n \in D$, $n \neq 0$, and an integer p such that $\exp(x, n) = V_A{}^p(\exp(x, n)) = \exp(x, A^{*p}n)$. We conclude $A^{*p}n = n$ and thus $K_p(A) \neq \{0\}$ which contradicts the hypothesis.

We observe that this is the case whenever A does not have eigenvalues which are roots of unity.

THEOREM 6. *If $S = (x_0, A)$ is an affine transformation of the m dimensional torus X and if there is an integer $p > 1$ such that $K_1(A) \neq K_p(A)$ then S is not ergodic.*

Proof. Because of Lemma 4 we need only construct a character which is periodic under V_S. We let

$$B_p = A^{p-1} + A^{p-2} + \cdots + A + I.$$

According to the hypothesis there is an $n \in D$, $n \neq 0$ for which $A^{*p}n = n$ but $A^*n \neq n$. If $n' = (A^* - I)n$ then $n' \neq 0$. Since $A^{*p} - I = B_p^*(A^* - I)$ and $(A^{*p} - I)n = 0$ we see that $B_p^*n' = 0$. We observe

$$A^{*p}n' = A^{*p}(A^* - I)n = (A^* - I)A^{*p}n = (A^* - I)n = n'.$$

We conclude the theorem in the following manner

$$\begin{aligned} V_{S^p} \exp(x, n') &= \exp(B_p x_0, n') \exp(A^p x, n') \\ &= \exp(x_0, B_p^* n') \exp(x, A^{*p} n') \\ &= \exp(x, n'). \end{aligned}$$

In order to make the next theorem more concise we make the following definition which is a generalization of rational independence.

Definition 7. If $x = (x_1, x_2, \cdots, x_m) \in R^m$ and if $D_0 \subset D$ is a submodule we say that the numbers $x_1, \cdots, x_m$ are rationally independent over D_0 if whenever $n \in D_0$ and (x, n) is rational then $n = 0$. When there is no possibility of confusion we will say x is rationally independent over D_0.

If $D_0 = D$ we observe that the previous definition reduces to the classical definition of rational independence. It is easy to see that x is independent over D_0 if and only if whenever $(x, n) \equiv 0 \bmod 1$ and $n \in D_0$ then $n = 0$.

THEOREM 8. *Let $S = (x_0, A)$ be an affine transformation of the m dimensional torus X. Suppose that $\{0\} \neq K_1(A) = K_p(A)$ for all $p \geqq 1$. Under these conditions the following two statements are true:*

a) *S is ergodic if and only if x_0 is rationally independent over $K_1(A)$.*

b) *S is not weakly mixing.*

Proof of a). Suppose x_0 is rationally dependent over $K_1(A) = K_p(A)$. Then there is an $n \in K_1(A)$, $n \neq 0$, or which $(x_0, n) = q/p$. We observe

$$\begin{aligned} V_{S^p} \exp(x, n) &= \exp(B_p x_0, n) \exp(A^p x, n) \\ &= \exp(x_0, (A^{*p-1} + \cdots + A^* + I)n) \exp(x, A^{*p}n) \\ &= \exp(x_0, pn) \exp(x, n) = \exp p(x_0, n) \exp(x, n) \\ &= \exp(q) \exp(x, n) = \exp(x, n). \end{aligned}$$

It follows from Lemma 4 that S is not ergodic.

Suppose now that S is not ergodic. We wish to construct an $n \in K_1(A)$, $n \neq 0$, for which (x_0, n) is rational. Since S is not ergodic it follows from Theorem 4 that there is an $n \in L$, $n \neq 0$, and an integer p such that

$$\begin{aligned} V_S{}^p \exp(x, n) &= \lambda_p(\exp(x, n)) V_A{}^p \exp(x, n) \\ &= V_A{}^p \exp(x, n) = \exp(x, n). \end{aligned}$$

From this we get two pieces of information. First of all

$$\exp(x, n) = V_A{}^p \exp(x, n) = \exp(x, A^{*p} n)$$

from which we conclude $n = A^{*p} n$ or that $n \in K_p(A) = K_1(A)$.

Secondly we see

$$\begin{aligned} 1 &= \lambda_p(\exp(x_0, n)) = \exp(x_0, (A^{*p-1} + A^{*p} + \cdots + A^* + I) n) \\ &= \exp(x_0, pn). \end{aligned}$$

From this we conclude that

$$(x_0, pn) = p(x_0, n) \equiv 0 \mod 1$$

or that (x_0, n) is rational.

Proof of b). Since $K_1(A) \neq \{0\}$ there is an $n \neq 0$ for which $V_A \exp(x, n) = \exp(x, A^* n) = \exp(x, n)$. It follows from Theorem 2 that S is not weakly mixing.

When A is the identity Theorem 8 becomes the measure theoretic version of Kronecker's little theorem.

3. Spectrum of affine transformations of compact Abelian groups. In this section we wish to examine in more detail the spectrum of the operator V_S $(S = (x_0, A) \in \mathrm{Aff}(X))$ acting on the Hilbert space $L_2(X)$. Our first theorem is a careful statement of the results implicit in the preceding two sections.

THEOREM 9. *The space $L_2(X)$ is the direct sum of two orthogonal subspaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$ with the following properties:*

a) $V_S : \mathfrak{H}_i \to \mathfrak{H}_i$, $i = 1, 2$.

b) *Let $V_i = V_S \mid \mathfrak{H}_i$, $i = 1, 2$. The unitary operator V_1 has pure point spectrum.*

c) $\mathfrak{H}_2 = \sum_\alpha \mathfrak{H}_\alpha$ *where each $\mathfrak{H}_\alpha$ is separable and has an orthonormal basis $\{f_{i,\alpha} : i = 0, \pm 1, \pm 2, \pm 3, \cdots\}$. Moreover $V_2 f_{i,\alpha} = f_{i+1,\alpha}$ and thus V_2 has pure continuous spectrum.*

Proof. We let $\mathfrak{H}_1$ be the subspace spanned by the set of all characters $\xi \in \hat{X}$ which have a finite orbit under V_A, that is there is an integer n for which $V_A{}^n\xi = \xi$. We let $\mathfrak{H}_2$ be the subspace spanned by the set of all characters which have infinite orbits under V_A. We verify easily that a) holds because $V_S\xi = \lambda_1(\xi) V_A\xi$ where $|\lambda_1(\alpha)| = 1$.

The space $\mathfrak{H}_1$ decomposes into the direct sum of finite dimensional subspaces which are stable under V_S. These subspaces are formed by choosing any character ξ for which there is an integer n such that $V_A{}^n\xi = \xi$ and taking the space spanned by $\xi, V_A{}^2\xi, \cdots, V_A{}^{n-1}\xi$. On any such subspaces V_1 is a finite dimensional unitary operator and thus has pure spectrum. This remark concludes the proof of b).

The proof of c) is clear when we observe that the subspaces $\mathfrak{H}_\alpha$ are merely the spaces generated by the orbits of the characters generating $\mathfrak{H}_2$ (see [6], p. 53).

A slight further examination of $\mathfrak{H}_1$ is of interest. Let $\xi \in \hat{X}$ and $\xi \in \mathfrak{H}_1$. Also let $\mathfrak{H}_1(\xi)$ be the space spanned by $\xi, V_A\xi, V_A{}^2\xi, \cdots, V_A{}^{n-1}\xi$ where n is the period of ξ under V_A. We observe $V_S(V_A{}^j(\xi)) = a_j V_A{}^{j+1}(\xi)$ for $j = 0, 1, 2, \cdots, n-1$. If we let $D(\xi)$ be the determinant of V_1 restricted to $\mathfrak{H}_1(\xi)$ we see that $D(\xi) = (-1)^{n-1} a_0 \cdot a_1 \cdot a_2 \cdots \cdot a_{n-1}$. Moreover the characteristic polynomial for V_1 restricted to $\mathfrak{H}_1(\xi)$ is given by $(-1)^n(\lambda^n - D(\xi))$. Then the eigenvalues are n-th roots of $D(\xi)$. Combining this with Theorem 8 we obtain

COROLLARY 10. *S is ergodic if and only if for each $\xi \in \hat{X} \cap \mathfrak{H}_1$, $\xi \neq 1$ we have $D(\xi) \neq 1$.*

We examine still further the subspace $\mathfrak{H}_1$ and the operator V_1. Let $\hat{H} = \mathfrak{H}_1 \cap \hat{X}$ and let $\hat{H}^* = \{x : \xi(x) = 1 \text{ if } \xi \in H\}$. It is not difficult to verify that $\hat{H}$ is a subgroup of the discrete group $\hat{X}$. We let $H = X/\hat{H}^*$. It is known then that $\hat{H}$ is the character group of H. The set $\hat{H}$ is stable under V_A so it follows that $\hat{H}^*$ is mapped onto itself under A. Thus we make $S = (x_0, A)$ act on H in the following manner $S(x + \hat{H}^*) = x_0 + Ax + H^*$.

THEOREM 11. *The space $\mathfrak{H}_1 = L_2(H)$ and the action induced by S on $L_2(H)$ is the same as the action of V_1 on $\mathfrak{H}_1$. The action of S on X is ergodic if and only if the action of S on H is ergodic. If S is ergodic on X then there is an element $h_0 \in H$ such that $S(h) = h + h_0$ for all $h \in H$ and the subgroup generated by h_0 is dense in H.*

Proof. The first sentence of the theorem merely sums up the paragraph immediately before the theorem. If we look at the action of V_S on $L_2(X)$

we see that the eigenfunctions must all lie in $\mathfrak{H}_1$. Thus S is ergodic on X if and only if the dimension of the eigenvalue one space of V_1 acting on $\mathfrak{H}_1$ is one. Since the action induced by S on $L_2(H)$ is the same as the action induced by V_1 on $\mathfrak{H}_1$ the second sentence holds.

If S is ergodic on X then S acting on H is also ergodic. This action on H induces the same unitary action as V_1 on $\mathfrak{H}_1$. Since V_1 acting on $\mathfrak{H}_1$ has pure point spectrum the results of the last sentence of the theorem follow from standard theorems in ergodic theory. (See [6], p. 46 et. seq.)

4. Strict ergodicity and standard forms of affine transformations on the torus. If Ω is a compact metric space and T a homeomorphism of Ω we recall the following definition.

Definition 12. The homeomorphism T of the compact metric space Ω is said to be strictly ergodic if there is a unique T invariant Borel measure μ with $\mu(\Omega) = 1$.

We point out that this is commonly called "uniquely ergodic" but we follow the usage of Furstenberg [4], p. 574, since many of the references are to this paper.

Let μ be any T invariant measure on Ω and let $C(\Omega)$ be the set of all continuous complex valued functions on Ω.

Definition 13. A point $\omega \in \Omega$ is said to be generic for (Ω, T, μ) if for each f in $C(\Omega)$ we have

$$\lim_N 1/N \sum_{n=0}^{N-1} f(T^n\omega) = \int_\Omega f\, d\mu.$$

If T is ergodic then it is direct consequence of the ergodic theorem that almost all ω (with respect to μ) are generic points for (Ω, T, μ).

We state the following known theorem for later reference. For its proof see for instance [4], p. 575.

THEOREM 14. *If μ is a T invariant Borel measure on the compact metric space Ω then the following statements are equivalent:*

1. *T is strictly ergodic on Ω with invariant measure μ.*

2. *Every point of Ω is generic for (Ω, T, μ).*

3. $1/N \sum_{n=0}^{N-1} f(T^n\omega)$ *converges uniformly in ω to* $\int_\Omega f\, d\mu$ *for each $f \in C(\Omega)$.*

We consider two compact metric abelian groups X_1 and X_2. Let $A_1 \in \mathrm{Aut}(X_1)$ and let $B: X_1 \to X_2$ be a homomorphism.

Let $\gamma = (\gamma_1, \gamma_2) \in X_1 \times X_2$. Let $X = X_1 \times X_2$ and let $A \in \operatorname{Aut}(X)$ be given by $A(x) = A(x_1, x_2) = (A_1(x_1), x_2 + B(x_1))$. If μ_i are normalized Haar measure on X_i we know that μ_i is A_i invariant and we observe that if $\mu = \mu_1 \times \mu_2$ then μ is A invariant. To see this we need merely examine the integral of functions in $C(X)$ of the form $f(x) = f_1(x_1) f_2(x)$ where $f_i \in C(X_i)$.

$$\begin{aligned}\int f(A(x))\,d\mu(x) &= \int\int f_1(A_1(x_1)) f_2(x_2 + B(x_1))\,d\mu_1(x_1)\,d\mu_2(x_2)\\ &= \int f_1(A_1(x_1))\,[\int f_2(x_2 + B(x_1))\,d\mu_2\ (x_2)]\,d\mu_1(x_1)\\ &= \int f_1(A_1(x_1))\,[\int f_2(x_2)\,d\mu_2(x_2)]\,d\mu_1(x_1)\\ &= \int f_1(x_1)\,d\mu_1(x_1)\cdot\int f_2(x_2)\,d\mu_2(x_2)\\ &= \int f(x)\,d\mu(x).\end{aligned}$$

The idea of the following theorem is essentially contained in [4], p. 578, but for completeness we include its proof here.

THEOREM 15. *In terms of the preceding notation let* $S = (\gamma, A) \in \operatorname{Aff}(X)$. *The product measure* μ *is invariant under* S. *If* $S_1 = (\gamma_1, A_1) \in \operatorname{Aff}(X_1)$ *is strictly ergodic and if* S *is ergodic for* μ *on* X *then* S *is strictly ergodic on* X.

Proof. Since translation preserves the product measure μ it is clear from the previous discussion that μ is S invariant.

We first observe that if $x = (x_1, x_2)$ is generic for the product measure μ then for any $y \in X_2$, $(x_1, x_2 + y)$ is generic for μ. To see this fix $y \in X_2$ and let $T(x_1, x_2) = (x_1, x_2 + y)$

$$\begin{aligned}\lim_N 1/N \sum_{n=0}^{N-1} f(S^n(x_1, x_2 + y)) &= \lim_N 1/N \sum_{n=0}^{N-1} f\cdot T(S^n(x_1, x_2)\\ &= \int f\circ T\ (x_1, x_2)\,d\mu = \int\int f(x_1, x_2 + y)\,d\mu_1(x_1)\,d\mu_2(x_2)\\ &= \int f(x_1, x_2)\,d\mu = \int f\,d\mu.\end{aligned}$$

Thus if we let $\pi\colon X_1 \times X_2 \to X_1$ be the projection onto X_1 and if $E(\mu)$ is the set of generic points for (X, S, μ) then we see that $E(\mu) = \pi(E(\mu)) \times X_2$.

Suppose ν is an ergodic S invariant measure on X. We wish to show $\mu = \nu$. To do this we need only show that there is a point $x \in X$ which is generic for both (X, S, μ) and (X, S, ν) since this implies $\int f\,d\mu = \int f\,d\nu$ for all $f \in C(X)$. On the Borel subsets of X_1 define the measure ν_1 as follows $\nu_1(E_1) = \nu(E_1 \times X_2)$. The measure ν_1 is S_1 invariant and from the strict ergodicity of S_1 it follows that $\nu_1 = \mu_1$. Let $E(\nu)$ be the set of generic points of (X, S, ν). From the fact that S is ergodic and from the ergodic theorem we obtain $\mu(E(\mu)) = \nu(E(\nu)) = 1$. Since $E(\nu) \subset \pi(E(\nu)) \times X_2$ we also see that $1 \geqq \nu_1(\pi(E(\nu)) = \nu(\pi(E(\nu)) \times X_2) \geqq \nu(E(\nu)) = 1$. And similarly

$\mu_1(\pi(E(\mu)))=1$. Since $\mu_1=\nu_1$ we have $\pi(E(\mu))\cap\pi(E(\nu))\neq\emptyset$. The fact that $E(\mu)=\pi(E(\mu))\times X_2$ implies $E(\mu)\cap E(\nu)\neq\emptyset$ which concludes the proof.

For the remainder of this section we shall only consider affine transformations on an m dimensional torus. As was pointed out earlier the automorphisms of the m dimensional torus are in one to one correspondence with the automorphisms of the free abelian group on m generators. With this in mind we will be able to make use of the following theorem. (At this point I would like to acknowledge some helpful suggestions by G. Seligman.)

THEOREM 16. *Let G be a finitely generated free abelian group and let σ be an endomorphism of G. Under these conditions there exists a sequence of subgroups $G=G_0\supset G_1\supset G_2\cdots\supset G_k$, $k\geqq 0$, such that the inclusions are proper; $\sigma\colon G_i\to G_{i+1}$ for $i=0,1,\cdots,k-1$; $\sigma(G_k)\subset G_k$ and $G_k/\sigma(G_k)$ is either trivial or is a torsion group. Moreover there exist subsets $B_i\subset G_i$ with $B_0\supset B_1\supset\cdots\supset B_k$ such that each B_i is a finite set of independent generators of G_i.*

Proof. Let $L=Q\otimes_Z G$ where Q is the additive group of rational numbers and Z the integers. L is a vector space over Q and σ becomes a linear transformation on L if it is defined on decomposable tensors as follows $\sigma(r\otimes g)=r\otimes\sigma(g)$. Let $L_i=\sigma^i(L)$ for $i=0,1,2,\cdots$. There is a first k for which $L_k=L_{k+1}$. We then obtain the properly descending chain of vector spaces

$$L=L_0\supset L_1\supset L_2\supset\cdots\supset L_k.$$

If we let $G_i=L_i\cap G$ then we have the properly descending chain

$$G=G_0\supset G_1\supset G_2\supset\cdots\supset G_k.$$

We see that $\sigma(G_i)=\sigma(L_i\cap G)\subset\sigma(L_i)\cap\sigma(G)\subset L_{i+1}\cap G=G_{i+1}$ for $i=0,1,2,\cdots,k-1$. We wish to show that $G_k/\sigma(G_k)$ is a torsion group if it is not trivial. If it were not then there would be an element $v\in G_k$ for which $nv\notin\sigma(G_k)$ for all positive integers n. Since $\sigma(L_k)=L_k\supset G_k$ there is an element $w=\sum r_j\otimes g_j\in L_k$ for which $\sigma(w)=\sum r_j\otimes\sigma(g_j)=v$. Choose q so that qr_j is an integer for each j. Then $qw\in L_k\cap G=G_k$ and $\sigma(qw)=q\sigma(w)=qv\in\sigma(G_k)$ which is a contradiction.

We must now establish the existence of the sets $B_0\supset B_1\supset\cdots\supset B_k$. Let B_k be any independent set of generators for the finitely generated free group G_k. Suppose we have constructed $B_{i+1}\supset B_{i+2}\supset\cdots\supset B_k$, $0\leqq i<k$. Consider G_i/G_{i+1} and the homomorphism $G_i/G_{i+1}\to L_i/L_{i+1}$ given by $g+G_{i+1}\to g+L_{i+1}$. This mapping is one to one. For if $g\in G$ and $g\in L_{i+1}$ then

$g \in L_{t+1} \cap G = G_{t+1}$. Thus we see that G_t/G_{t+1} is torsion free and we have the diagram.

$$0 \to G_{t+1} \to G_t \to G_t/G_{t+1} \to 0.$$

Since G_t/G_{t+1} is a finitely generated free group G_t is the direct sum of G_{t+1} and another group. Thus the basis B_{t+1} of G_{t+1} may be extended to a basis B_t of G_t. Proceeding inductively in this manner we reach B_0 and the conclusion of the proof.

We have thus shown that for any endomorphism σ of a finitely generated free abelian group G there is a basis of G with respect to which σ has the following form

$$\left(\begin{array}{c|c} \alpha & 0 \\ \hline * & \tau \end{array}\right)$$

Where τ is a square matrix, α is a triangular block

$$\left|\begin{array}{cccc} 0 & & & \\ & 0 & & \\ & & \ddots & \\ X & & & 0, \end{array}\right.$$

and there may be non zero entries in the section indicated with a $*$. We will refer to this as a standard form for σ and a basis as the one in Theorem 16 shall be called a standard basis with respect to σ. In the case where σ is nilpotent ($\sigma^k = 0$) then in Theorem 16 $G_k = L_k = \{0\}$, and the block τ disappears from the standard matrix form and σ has a triangular form.

We will now examine in more detail the structure of an affine transformation $S = (x_0, A)$ of a finite dimensional torus X. We will restrict ourselves only to the case where S is ergodic. The results of Section 2 state that either $\{0\} = K_p(A)$ for all $p \geqq 1$ or $\{0\} \neq K_1(A) = K_p(A)$ for all $p \geqq 1$. In other words either A has no eigenvalues which are roots of unity or A has some eigenvalues equal to 1 but no eigenvalues equal to any other root of unity. We recall that A^* is an automorphism of the finitely generated free group $\hat{X}$ and state the following obvious remark

Remark 17. *If* $K_p(A) = \{0\}$ *for all* $p \geqq 1$ *then the number* k *for* $\sigma = A^* - I$ *in Theorem* 16 *is equal to zero and the standard form of* σ *has no triangular part.*

We know that S is strongly mixing and if $K_1(A) = \{0\}$, then $A - I$ is non-singular and there is a point $x \in \boldsymbol{R}^m$ with $Ax - x = -x_0$, $Ax + x_0 = x$.

Reducing modulo 1 we find that S has a fixed point in the torus and so it cannot be strictly ergodic.

We now wish to examine the case where $\{0\} \neq K_1(A) = K_p(A)$, $p \geqq 1$.

THEOREM 18. *If S is ergodic and if $\{0\} \neq K_1(A) = K_p(A)$ and if $A - I$ is nilpotent then A has a standard form*

$$\begin{pmatrix} 1 & & & & * \\ & 1 & & & \\ & & 1 & & \\ 0 & & & 1 & \\ & & & & 1 \end{pmatrix} \quad \text{and} \quad x_0 = \begin{pmatrix} \xi_1 \\ \cdot \\ \cdot \\ \cdot \\ \xi_n \end{pmatrix}$$

and at least ξ_n is irrational. Moreover S is strictly ergodic.

Proof. If $A^* - I$ is nilpotent then Theorem 16 implies that there is a basis $e_1, e_2, \cdots, e_n$ such that

$$\begin{aligned} A^*e_1 &= e_1 \\ A^*e_2 &= a_{21}e_1 + e_2 \\ A^*e_3 &= a_{31}e_1 + a_{32}e_2 + e_3 \\ A^*e_n &= a_{n1}e_1 + a_{n2}e_2 + \cdots + a_{n,n-1}e_{n-1} + e_n. \end{aligned}$$

With respect to this basis A^* has the form

$$\begin{pmatrix} 1 & & & & \\ & 1 & & & 0 \\ & & 1 & & \\ * & & & 1 & \\ & & & & 1 \end{pmatrix}$$

and thus A has the super triangular form. The fact that ξ_n is irrational follows immediately from the fact that S is ergodic and thus by Theorem 8 x_0 is rationally independent over $K_1(A)$.

In order to obtain strict ergodicity of S we first observe that translation by an irrational of the reals mod 1 is strictly ergodic. We use the ergodicity of S and Theorem 15 to conclude that S is strictly ergodic.

The affine transformations $S = (x_0, A)$ also have another interesting spectral property when $A - I$ is nilpotent. We recall the following definition of quasi-eigenfunctions (see [6], p. 57). Let $\mathfrak{G}_0 \subset L_2(X)$ be the set of all constant functions. For $i > 0$ let $\mathfrak{G}_i \subset L_2(X)$ be the set of all functions f such that $V_S f = g \cdot f$ where $g \in \mathfrak{G}_{i-1}$. The set $\mathfrak{G}_i$ is the set of quasi-eigenfunctions of degree i. S is said to have quasi-discrete spectrum if $\bigcup_{i=0}^{\infty} \mathfrak{G}_i$ spans $L_2(X)$.

THEOREM 19. *If $S=(x_0,A)$ and if $A-I$ is nilpotent then S has quasi-discrete spectrum.*

Proof. From the previous theorem we know the S may be written as follows

$$\begin{aligned} &x_1\to\xi_1+x_1\\ &x_2\to\xi_2+a_{21}x_1+x_2\\ &x_3\to\xi_3+a_{31}x_1+a_{32}x_2+x_3\\ &x_n\to\xi_n+a_{n1}x_1+a_{n2}x_2+a_{n3}x_3+\cdots+a_{n,n-1}x_{n-1}+x_n. \end{aligned}$$

Where the ξ_i are fixed real numbers, the a_{ij} are fixed integers and the x_i are real numbers. The transformation is always taken mod-1.

Let $f_0(x)\equiv 1$, $f(x,p_1,p_2\cdots p_j)=\exp(\sum_{i=1}^{j}p_ix_i)$, $j=1,2,\cdots,n$, and the p_i are arbitrary integers. If $j=1$ the functions $f(x,p_1)$ are quasi-eigenfunctions of degree one. Suppose we have shown that if $j=k$ the functions $f(x,p_1,\cdots,p_k)$ are quasi-eigenfunctions of degree k. We observe that

$$V_Sf(x,p_1,\cdots,p_{k+1})=cf(x,p'_1,\cdots,p'_k)f(x,p_1,\cdots,p_{k+1}).$$

Since a constant times a quasi-eigenfunction is a quasi-eigenfunction of the same degree it follows that $f(x,p_1,\cdots,p_{k+1})$ is a quasi-eigenfunction of degree $k+1$. The fact that f_0 together with all the functions $f(x,p_1,\cdots,p_j)$, $j=1,2,\cdots,n$ span $L_2(X)$ shows that S has quasi-discrete spectrum.

The case in which $A-I$ is not nilpotent does not allow us to make the assertions of the preceding two theorems. For instance if we let $S_1=(x_0,A_1)$ be ergodic with A_1-I nilpotent and if we let $S_2=A_2$ be a strongly mixing automorphism for which some ergodic averages fail to exist [4], p. 586 then $S_1\times S_2$ is ergodic but not strictly ergodic since some ergodic averages will fail to exist.

Example 20. We wish to give an example of an affine transformation $S=(x_0,A)$ on the three dimensional torus which is ergodic, not strongly mixing but does not have quasi-discrete spectrum. Let A have the matrix

$$\begin{pmatrix}1&0&0\\0&3&2\\0&2&1\end{pmatrix}\text{ and let } x_0=\begin{pmatrix}\gamma\\0\\0\end{pmatrix}$$

where γ is irrational. Under these conditions S is ergodic but not strongly mixing. We also see that $A-I$ is not nilpotent. The transformation S is given by

$$\begin{aligned} &x_1\to x_1+\gamma\\ &x_2\to 3x_2+2x_3\\ &x_3\to 2x_2+x_3. \end{aligned}$$

Let the elements of the form $(0, x_2, x_3)$ be denoted by Y and those of the form $(x_1, 0, 0)$ be denoted by X. Let $\tau\colon Y \to Y$ be given by the matrix $\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$. Then S is given by

$$x \to x + \gamma$$
$$y \to \tau(y)$$

and τ is strongly mixing on Y. The set $\mathfrak{G}_1$ of quasi-eigenfunctions of degree one are the functions of the form $f(x,y) = c\exp(nx)$ where n is any integer and c is any complex number. If g is in $\mathfrak{G}_2$ then $V_S g = f \cdot g$. Now sinve V_S is unitary we have

$$(g,g) = (V_S g, V_S g) = (fg, fg) = (c\exp(nx) g(x,y), c\exp(nx) g(x,y)$$
$$= c\bar{c}(g,g).$$

Thus we have $|c| = 1$ and the quasi-eigenvalue f has modulus 1. We wish to show that $\mathfrak{G}_2 = \mathfrak{G}_1$ and consequently $\mathfrak{G}_p = \mathfrak{G}_1$ for $p \geqq 1$ Since $\mathfrak{G}_1$ does not span $L_2(X \times Y)$ this would show that S does not have quasi-discrete spectrum. Suppose $g \in \mathfrak{G}_2$ then

$$g(x,y) = \sum_{\xi} a_\xi(x)\xi(y)$$

where the sum is taken over $\hat{Y}$ and

$$a_\xi(x) = \int_Y g(x,y)\xi(y)\,dy.$$

Thus $V_S g(x,y) = \sum a_\xi(x+\gamma)\xi(\tau(y))$. Since $g \in \mathfrak{G}^2$ we have

$$V_S g(x,y) = f(x) g(x,y) = \sum_{\xi} a_\xi(x) f(x)\xi(\gamma).$$

Comparing coefficients we obtain

$$a_\xi(x+\gamma) = a_{\tau^*(\xi)}(x) f(x).$$

Since $|f(x)| = 1$ and τ^* has all infinite orbits on $\hat{Y}$ except in the case where $\xi \equiv 1$ and $\sum |a_\xi|^2 < +\infty$ we see that $a_\xi = 0$ if $\xi \not\equiv 1$. This shows that

$$g(x,y) = a_1(x).$$

But now it is easy to see that $a_1(x)$ is in $\mathfrak{G}_1$ since we need only examine the rotation $x \to x + \gamma$. Thus $\mathfrak{G}_1 = \mathfrak{G}_2$.

5. The theorem of H. Weyl. In this section we will apply the results of Sections 2 and 4 to prove a theorem of Hermann Weyl on uniform distribution [8]. We say that a sequence of numbers x_n is uniformly distributed

mod 1 if for each continuous complex valued function f, which also has period 1, we have

$$\lim_N 1/N \sum_{n=0}^{N-1} f(x_n) = \int_0^1 f(x)\,dx.$$

Theorem 21 (H. Weyl). Let $q(z) = a_m z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0$ be a real polynomial with $a_m \neq 0$ and $m > 1$. If there is an $i \neq 0$ for which a_i is irrational then the sequence $q(n)$, $n = 0, 1, 2, \cdots$, is uniformly distributed mod 1.

It is not difficult to see that it suffices to prove the theorem for the case where a_m is irrational. Suppose the theorem were true whenever the leading coefficient of q is irrational. Let

$$q(x) = a_{m+j}x^{m+j} + \cdots + a_j x^j + \cdots + a_0$$

and suppose a_j is the first irrational coefficient. Choose M such that Ma_{m+j}, $Ma_{m+j-1}, \cdots, Ma_{j+1}$ are integers. Each integer n may be written $n = kM + s$, $s = 0, 1, \cdots, M-1$. We wish to show that $q(kM+s)$ is uniformly distributed in k for each integer $s = 0, 1, 2, \cdots, M-1$. This follows from the fact that $q(kM+s)$ is congruent mod 1 to a polynomial of degree j whose leading coefficient is irrational. Now it is not difficult to see that $q(n)$ is uniformly distributed.

We will prove this reduced form of the theorem. In order to prove this theorem we will construct a strictly ergodic affine transformation of the $m+1$ dimensional torus one of whose coordinate sequence is given by an arbitrary preassigned polynomial of degree m with leading coefficient irrational. Let

$$\xi_0 = \begin{pmatrix} \gamma \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and let } A = \begin{pmatrix} 1 & & & 0 \\ 1 & \cdot & & \\ & \cdot & \cdot & \\ 0 & & 1 & 1 \end{pmatrix}$$

and let $S = (\xi_0, A)$. That is S is the transformation of the $m+1$ dimensional torus given by

$$\begin{aligned} x_0 &\to x_0 + \gamma \\ x_1 &\to x_1 + x_0 \\ x_2 &\to x_2 + x_1 \\ x_m &\to x_2 + x_1. \end{aligned}$$

The reader should be warned that for simplicity of expression I have taken A in sub diagonal form where as in the previous theorems it was in super diagonal form. This of course does not alter the arguments. We must

study the form of the powers of A and we give our results in the next two lemmas.

LEMMA 21.1. *If* $A^n = (C^n{}_{ij})$ *for* $n = 0, 1, 2, \cdots$ *then* A^n *is symmetric about its secondary diagonal. That is* $C^n{}_{m-j,m-i} = C^n{}_{i,j}$.

Proof. The statement is true if $n = 0$. Suppose it is true for $n - 1$. Since $A^n = A \cdot A^{n-1} = A^{n-1}A$ we obtain

$$A^n_{i,j} = C^{n-1}_{i-1,j} + C^{n-1}_{i,j} = C^{n-1}_{i,j} + C^{n-1}_{i,j+1}\,.$$

Applying this formula and the induction hypothesis we obtain

$$\begin{aligned} C^n_{m-j,m-i} &= C^{n-1}_{m-(j+1),m-i} + C^{n-1}_{m-j,m-i} \\ &= C^{n-1}_{i,j+1} + C^{n-1}_{i,j} = C^n_{i,j}\,. \end{aligned}$$

For our purposes we must also compute the entries of the first column of A^n. We consider the infinite array

$$\begin{array}{cccccccc} 1 & 1 & 1 & 1 & 1 & 1 & \cdot & \cdot\ \cdot \\ 0 & 1 & 2 & 3 & 4 & 5 & \cdot & \cdot\ \cdot \\ 0 & 0 & 1 & 3 & 6 & 10 & \cdot & \cdot\ \cdot \\ 0 & 0 & 0 & 1 & 4 & 10 & \cdot & \cdot\ \cdot \\ 0 & 0 & 0 & 0 & 1 & 5 & \cdot & \cdot\ \cdot \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdot & \cdot\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{array}$$

whose law of formation is $p_0(j) = 1$ for $j \geqq 0$, $p_i(0) = 0$ for $i > 0$ and $p_i(j) = p_i(j-1) + p_{i-1}(j-1)$ for $i > 0$ and $j > 0$. Where i is the row index and j the column index and $i, j = 0, 1, 2, \cdots$. Observing that this array is merely Pascal's triangle we also obtain

$$p_0(j) = 1, \quad j = 0, 1, 2, \cdots, \quad i = 1, 2, 3, \cdots$$

$$p_i(j) = \frac{j(j-1)\cdots(j-i+1)}{i!}, \quad j = 0, 1, 2, 3 \cdots, i = 1, 2, 3, \cdots.$$

Consequently in all cases we see that $p_i(j)$ is a polynomial in j with rational coefficients and degree i. Another fact that we will use later is that

$$(*) \qquad \sum_{k=0}^{j-1} p_i(k) = p_{i+1}(j) \text{ for } \begin{matrix} j = 1, 2, 3, \cdots \\ i = 0, 1, 2, \cdots. \end{matrix}$$

LEMMA 21.2. *If* $A^n = (C^n{}_{ij})$ *then* $C^n{}_{i0} = p_i(n)$ *for* $i = 0, 1, 2, \cdots, m$ *and* $n = 0, 1, 2, \cdots$.

Proof. The statement is true if $n=0$ and the induction follows immediately from the two formulae

$$C^{n+1}{}_{i,0} = C^{n}{}_{i-1,0} + C^{n}{}_{i,0}$$
$$p_i(n+1) = p_{i-1}(n) + p_i(n).$$

We now prove Theorem 21. Let $q(z) = a_{m+1}z^{m+1} + a_m z^m + \cdots + a_1 z + a_0$ where a_m is irrational and $m > 0$. Consider the transformation $S = (\xi_0, A)$. For any $x = (x_0, x_1 \cdots, x_m)$ we have

$$S^n(x) = \xi_0 + A\xi_0 + A^2\xi_0 + \cdots + A^{n-1}\xi_0 + A^n x.$$

If $y^n = (y^n{}_0, y^n{}_1, \cdots, y^n{}_m) = S^n(x)$ we wish to compute $y^n{}_m$. From Lemmas 21.1 and 21.2 and the formula (*) we see that

$$\begin{aligned} y^n{}_m &= \gamma(p_m(0) + p_m(1) + p_m(2) + \cdots + p_m(n-1)) \\ &\quad + x_0 p_m(n) + x_1 p_{m-1}(n) + \cdots + x_m p_0(n) \\ &= \gamma p_{m+1}(n) + x_0 p_m(n) + x_1 p_{m-1}(n) + \cdots + x_m p_0(n). \end{aligned}$$

Since the degree in n of each polynomial p_j is j and since each polynomial has rational coefficients there is an irrational number γ and real numbers $x_0, x_1, \cdots, x_m$ for which

$$q(n) = y^n{}_m.$$

Now it follows from Theorem 15 that $S = (\xi_0, A)$ is strictly ergodic. This implies uniform distribution of the points $S^n x$ for each x. In particular if f is a continuous function on $m+1$ space with period 1 in each variable we have

$$\lim_N 1/N \sum_{n=0}^{N-1} f(S^n x) = \int_0^1 \cdots \int_0^1 f(x_0, \cdots, x_m)\, dx_0, \cdots, dx_m.$$

If g is a continuous function of a single variable and has period one we let $f(x) = g(x_m)$. We then have

$$\lim_N 1/N \sum_{n=0}^{N-1} g(y^n{}_m) = \lim_N 1/N \sum_{n=0}^{N-1} g(q(n)) = \int_0^1 g(z)\, dz$$

which concludes the theorem.

I would like to point out that all the affine transformations $S = (x_0, A)$ where A has diagonal form may be imbedded in the nilflows of [3]. In fact

in Chapter VIII of [3] L. Auslander and the author used a certain nilflow to prove a theorem similar to the Hermann Weyl theorem. Our model there was different from the one used here and we obtained a slightly weaker theorem.

YALE UNIVERSITY.

REFERENCES.

[1] L. M. Abramov, "Metric automorphisms with quasi-discrete spectrum," *Izvestiya Akademii Nauk SSSR*, Vol. 26 (1962), pp. 513-530.

[2] H. Anzai, "Ergodic skew product transformations on the torus," *Osaka Mathematical Journal*, vol. 3 (1951), pp. 83-99.

[3] L. Auslander, L. Green, and F. Hahn, "Flows on homogeneous spaces," *Annals of Mathematics Studies*, No. 53, Princeton, N. J.

[4] H. Furstenberg, "Strict ergodicity and transformations of the torus," *American Journal of Mathematics*, vol. 83 (1961), pp. 573-601.

[5] P. R. Halmos, "On automorphisms of compact groups," *Bulletin of the American Mathematical Society*, vol. 49 (1943), pp. 619-624.

[6] ———, *Lectures on Ergodic Theory*, The Mathematical Society of Japan (1956).

[7] E. Hopf, *Ergodentheorie*, Chelsea, N. Y. (1948).

[8] H. Weyl, "Über die Gleichverteilung von Zahlen mod. Eins," *Mathematische Annalen*, vol. 77 (1916), pp. 313-352.

FINITE RINGS HAVING A CYCLIC MULTIPLICATIVE GROUP OF UNITS.*

By Robert W. Gilmer, Jr.[1]

This paper determines all finite commutative rings R with identity such that the multiplicative group of units of R is cyclic. The method of proof is straightforward, though the details sometimes become involved. We easily reduce the general case to that of determining all finite primary rings which have a cyclic multiplicative group of units. If S is such a primary ring, if g generates the multiplicative group of units of S, and if S_0 is the subring of S generated by the identity element e of S, then we can show that $S=S_0[g]$. Therefore S is a homomorphic image of $S_0[X]$ and of $Z[X]$. In the case when $S_0 \cong GF(p)$ for some prime p, $S_0[X]$ is a principal ideal domain whose ideal structure is well known. If S_0 is not a field, we consider S as a homomorphic image of $Z[X]$. The final result obtained is:

Each of the following classes consists of finite primary rings R having a cyclic multiplicative group of units. Any finite primary ring with a cyclic multiplicative group of units is isomorphic to an element of one and only one of these classes.

(A) $GF(p^k)$.

(B) $Z/(p^m)$ *where p is an odd prime and* $m>1$.

(C) $Z/(4)$.

(D) $\Pi_p[X]/(X^2)$ *where p is prime and* $\Pi_p \cong GF(p)$.

(E) $\Pi_2[X]/(X^3)$.

(F) $Z[X]/(4, 2X, X^2-2)$.

We shall use $|A|$ to denote the cardinality of a finite set A. If R is a ring with identity, G_R will denote the multiplicative group of units of R.

THEOREM 1. *If R is a finite commutative ring with identity, then R is a direct sum of primary rings $R_1, \cdots, R_k$ and G_R is a direct product of*

* Received July 19, 1963.

[1] This research was supported by the Office of Naval Research, Grant NONR(G) 00099-62. The author wishes to thank Professor C. W. Curtis for suggesting this problem to him.

$G_{R_1}, \cdots, G_{R_k}$. *G_R is cyclic if and only if each G_{R_i} is cyclic and* $(|G_{R_i}|, |G_{R_j}|) = 1$ *for* $1 \leqq i < j \leqq k$.

Proof. Since R is finite, R satisfies the descending chain condition and is therefore a direct sum of primary rings. [4; p. 205, Theorem 3]. If $r = \sum r_i$ where $r_i \in R_i$, we may easily show that r is a unit in R if and only if r_i is a unit of R_i for all i. Thus $G_R \cong G_{R_1} \otimes \cdots \otimes G_{R_k}$. The last assertion follows from elementary group theory.

Theorem 1 reduces the problem of finding all finite commutative rings R with identity such that G_R is cyclic to the problem of determining all finite primary rings with this property.

LEMMA 1. *If R is a finite commutative ring with identity such that G_R is cyclic and if f is a homomorphism from R onto a ring S, then f induces a homomorphism from G_R onto G_S. In particular, G_S is cyclic.*

Proof. We first suppose that R is a primary ring with maximal ideal M. In this case $G_R = R - M = \{g^i\}_{i=1}^n$ for some $g \in G_R$. If $S = (0)$, the lemma is true. If $S \neq (0)$, the kernel A of f is contained in M. If $m \in M$, then $f(m) \in f(M) \subset S$ so that $f(m)$ is not a unit of S. Therefore if $s = f(v) \in G_S$, then $v \in G_R$ so that $v = g^i$ and $s = [f(g)]^i$ for some integer i. The assertion then follows in case R is a primary ring.

In the general case $R = R_1 \oplus \cdots \oplus R_n$ is a direct sum of primary rings. We suppose $g = \sum g_i$ where $g_i \in R_i$ and g generates G_R. Now

$$|G_R| = |(g)| = |(g_1)| \cdots |(g_n)| = |G_{R_1}| \cdots |G_{R_n}|.$$

Since $|(g_i)| \leqq |G_{R_i}|$ for each i, $|(g_i)| = |G_{R_i}|$ for each i so that g_i generates G_{R_i}. If A is the kernel of f then $A = A_1 \oplus \cdots \oplus A_n$ where $A_i = AR_i$ is an ideal of R_i. Also $S \cong R/A \cong (R_1/A_1) \oplus \cdots \oplus (R_n/A_n)$. By the preceding case G_{R_i/A_i} is cyclic and is generated by $\bar{g}_i$. Further, $|G_{R_i/A_i}|$ is a divisor of $|G_{R_i}|$ so that $(|G_{R_i/A_i}|, |G_{R_j/A_j}|) = 1$ for $1 \leqq i < j \leqq n$. Theorem 1 now implies that G_S is cyclic and is generated by $\sum f(g_i) = f(\sum g_i) = f(g)$. This completes the proof.

THEOREM 2. *Let S be a finite primary ring with maximal ideal M. For some prime p, $|S| = p^s$. If S_0 is the subring of S generated by the identity element e of S, and if g generates G_S, then $S = S_0[g]$. Hence S is a homomorphic image of $S_0[X]$ and therefore also a homomorphic image of $Z[X]$.*

Proof. S/M is a finite field of characteristic p for some prime p. Thus

$pe \in M$ and $(pe)^k = p^k e = 0$ for some $k \in Z$. This implies $p^k a = 0$ for all $a \in S$ so that considered as an additive group, S is a p-group and $|S| = p^s$ for some s. [1, p. 43].

Evidently $S_0[g]$ contains every unit of S. If $m \in M$, then $m + g \notin M$ so that $m + g = g^i$ for some i and $m = g^i - g \in S_0[g]$. Therefore $S = S_0[g]$ as asserted. The last part of the theorem is obvious.

Theorem 2 greatly simplifies the problem under consideration for it shows we need consider only residue class rings of $S_0[X]$. An important special case is when $S_0 \cong \Pi_p$ for then $\Pi_p[X]$ is a principal ideal domain. Theorem 3 determines all finite primary rings of characteristic p which have a cyclic multiplicative group of units.

THEOREM 3. *Suppose R is a finite primary ring of characteristic p and with maximal ideal M. Let $|M| = p^r$ and $|R| = p^{r+k}$. Suppose further that g generates G_R and the homomorphism $f: \Pi_p[X] \to R$ is defined by $f(\sum a_i X^i) = \sum a_i g^i$. The kernel Q of f is a primary ideal and $\sqrt{Q} = P$ is maximal. In fact, $P = (q(X))$ for some irreducible polynomial $q(x)$ of degree k, and $Q = (q^n(X))$ for some integer n.*

The following statements hold:

a) *if $n \geqq 2$, then $k = 1$.*

b) *$n \leqq 3$ and if $n = 3$, $p = 2$. Consequently, it is necessary that either R is a finite field, $R \cong \Pi_p[X]/(X^2)$, or $p = 2$ and $R \cong \Pi_2[X]/(X^3)$. This condition is also sufficient in order that R be a primary ring of characteristic p for some prime p and that G_R be cyclic.*

Proof. By Theorem 2, f is onto so that $R \cong \Pi_p[X]/Q$. Thus Q is primary and P is maximal because R is a primary ring. That $P = (q(X))$ and $Q = (q^n(X))$ for some n then follow because $\Pi_p[X]$ is a principal ideal domain. We have $R/M \cong \Pi_p[X]/P$ and $|R/M| = p^k$. But $|\Pi_p[X]/P| = p^s$ where s is the degree of $q(X)$. Hence $s = k$.

If $n = 1$, R is a finite field. Now suppose $n \geqq 2$. If $a, b \in \Pi_p$, then $[1 + (a + bX) q^{n-1}(X)]^p \equiv 1 \pmod Q$. For $k \geqq 2$, this gives us $p^2 - 1$ units of R of order p. Hence $k = 1$. Without loss of generality we can take $q(X) = X$. Now if $(n-2)p \geqq n$, then $[1 + (a + bX)X^{n-2}]^p \equiv 1 \pmod Q$, which again gives us $p^2 - 1$ units of R of order p. Therefore $n > p(n-2) \geqq 2(n-2)$ so that $n < 4$ and if $n = 3$, $p = 2$. This proves statements a) and b) of the theorem.

It is well known that G_R is cyclic if R is a finite field. If $R = \Pi_2[X]/(X^3)$, $|G_R| = 4$ and $\overline{1+X}$ has order 4. If $R = \Pi_p[X]/(X^2)$, $\overline{1+X}$ has order p and for some element t of Π_p, $\bar{t}$ has order $p-1$. Hence $\overline{t(1+X)}$ has order $p(p-1) = |G_R|$.

We now turn to the case of a primary ring R of characteristic p^m where $m > 1$. In this case it will be important to consider R as a homomorphic image of $Z[X]$.

THEOREM 4. *Let R be a finite primary ring of characteristic p^m, $m > 1$, such that g generates G_R. If p is an odd prime $R \cong Z(p^m)$. If $p = 2$, then either $R \cong Z/(4)$ or $R \cong Z[X]/(4, 2X, X^2 - 2)$.*

Proof. Let $\mu: Z[X] \to R$ be the homomorphism defined by $\mu(\sum a_i X^i) = \sum a_i g^i$. The kernel Q of μ is a primary ideal containing p^m and its radical P is a maximal ideal containing p. We have

$$Z/p^m Z = Z/(Q \cap Z) \cong (Q + Z)/Q = S \subseteq Z[X]/Q \cong R.$$

We identify R and $Z[X]/Q$. S is a primary ring and G_S is the unique subgroup of G_R having order $p^{m-1}(p-1)$. In particular, every element of G_R having order a divisor of $p^{m-1}(p-1)$ is determined by an element of Z. Moreover, G_S is cyclic so that p is odd or $p = m = 2$. [3; pp. 114-5, Exercises 4-6].

We next wish to show $px - q \in Q$ for some $q \in Z$. This will greatly simplify the process of determining a complete set of residues of Q in $Z[X]$. We consider $b = (pX + 1)^{p^{m-1}}$. We shall show $b \equiv 1 \pmod{Q}$. To do so it suffices to show all coefficients of $b - 1$ are divisible by p^m. This follows immediately by induction on m. Hence $pX + 1$ has order $p^r \bmod Q$ for some $r \leq m - 1$. As shown previously, this implies $pX + 1 \equiv z \pmod{Q}$ or $pX - q \in Q$ for some $q \in Z$. Now $pX - q \in P$ implies $q = pc$ for some integer c. Since $p(X - c) \in Q$ and $p \notin Q$, $X - c \in P$. It follows that $A = (p, X - c) \subseteq P \subseteq Z[X]$. A, however, is maximal in $Z[X]$ so that $P = (p, X - c)$. Further, as we have just shown, $p(X - c) \in Q$.

If $q(X)$ is a monic polynomial of minimal degree in Q and if $q(X)$ has degree s, then modulo Q, every element of $Z[X]$ may be reduced to the form $\sum_{i=0}^{s-1} a_i X^i$ where $\{a_{s-1}, \cdots, a_1\} \subseteq \{0, 1, \cdots, p-1\}$ and $0 \leq a_0 \leq p^m - 1$. (Noting that $pX^w \equiv pc^w(Q)$.) Further if $a = \sum a_i X^i \in Q$, then each $a_i = 0$ for if $a_i \neq 0$ for some $i > 0$, we would obtain a contradiction to the mini-

mality of s since a_i is then a unit modulo Q. Then $a = a_0 \in Q$ so that $a = 0$ also. Therefore $|Z[X]/Q| = p^{m+s-1}$ in this case. Further if

$$\sum_{i=0}^{s-1} a_i X^i - p(\sum_{i=0}^{s-1} b_i X^i) - \overset{s-1}{\underset{i=0}{b_i}} X^i \mathrm{Q} = \sum (a_i - pb_i) X^i \in Q,$$

then $a_i = 0$ for $i > 0$ and p divides a_0. This observation shows that the set of all elements $\sum_{i=0}^{s-1} c_i X^i$ where $\{c_i\} \subseteq \{0, \cdots, p-1\}$ is a complete set of residues of the ideal $Q + (p)$ in $Z[X]$. Thus $Z[X]/Q + (p)$, being a homomorphic image of $Z[X]/Q$, is a primary ring of characteristic p, has order p^s, and has a cyclic multiplicative group of units by Lemma 1. By Theorem 3, $Q + (p) = (p, (X-c)^n)$ for some integer n. Since

$$p^s = p^n = |Z[X]/(p, (X-c)^n)|,$$

$n = s$ so that $(X-c)^s = a(X) + pb(X)$ for some $a(X) \in Q$, $b(X) \in Z[X]$. Therefore, $(X-c)^s \equiv py(Q)$ for some $y \in Z$. We have now shown that

$$(p^m, p(X-c), (X-c)^s - py) = B \subseteq Q.$$

However, it is clear that $|Z[X]/B| \leqq p^{m+s-1}$ so that $Q = B$. We note also that $(X-c)[(X-c)^s - py] = (X-c)^{s+1} - yp(X-c) \equiv 0(Q)$. Thus $(X-c)^{s+1} \equiv 0(Q)$.

We have already seen that if $s = 1$, $R \cong Z/(p^m)$. We now assume that $s > 1$. We then observe that

$$X^{p^{m+s-2}} \equiv [(X-c)+c]^{p^{m+s-2}} \equiv (X-c)^{p^{m+s-2}} + c^{p^{m+s-2}} \equiv c^{p^{m+s-2}} \pmod{Q}$$

because $p^{m+s-2} \geqq s+1$ and $p(X-c) \equiv 0 \pmod{Q}$. Further

$X^{p^{m+s-3}} \equiv c^{p^{m+s-3}}$ if $p^{m+s-3} \geqq s+1$. Now $p^{m+s-3} \geqq p^{m-1}$ so that $p^{m-1}(p-1)$

divides $p^{m+s-3}(p-1)$. Because c determines a unit of $Z/(p^m)$ (Note $c \equiv X \not\equiv 0 \pmod{p}$.),

$c^{p^{m-1}(p-1)} \equiv 1 \pmod{Q}$. Thus

$$c^{p^{m+s-3}(p-1)} \equiv 1 \pmod{Q} \text{ or } c^{p^{m+s-2}} \equiv c^{p^{m+s-3}} \pmod{Q}.$$

It follows that $X^{p^{m+s-2}} \equiv X^{p^{m+s-3}} \pmod{Q}$ or

$$X^{p^{m+s-3}(p-1)} - 1 \equiv 0 \pmod{Q} \text{ if } p^{m+s-3} \geqq s+1.$$

However, P has index p in $Z[X]$ so that the multiplicative group of units of $Z[X]/Q$ has order $p^{m+s-2}(p-1)$. Since g generates G_R we conclude

that $p^{m+s-3} < s+1$. But for p prime and m and $s \geqq 2$, this is possible only if $p = m = s = 2$.

Therefore, either $R \cong Z[X]/(4, 2X, X^2 - 2)$ or $R \cong Z[X]/(X^2, 2X, 4)$. One can verify directly that the multiplicative group of units of the first of these two rings is cyclic, while that of the second ring is not.

Remarks. 1. A finite ring R with identity such that G_R is cyclic need not be commutative. We may take R to be the collection of 2×2 upper triangular matrices with entries from Π_2.

2. If R is a finite commutative ring with identity such that g generates G_R and if R_0 is the subring of R generated by the identity e of R, it need not be true that $R = R_0[g]$. That is, not every element of R need be a sum of units. $R = \Pi_2 \oplus \Pi_2$ is such a ring.

UNIVERSITY OF WISCONSIN.

REFERENCES.

[1] Marshall Hall, *Theory of Groups*, Macmillan (New York), 1959.
[2] B. L. van der Waerden, *Modern Algebra*, vol. 1, Ungar (New York), 1953.
[3] ———, *Modern Algebra*, vol. 2, Ungar (New York), 1950.
[4] O. Zariski and P. Samuel, *Commutative Algebra*, vol. 1, Van Nostrand (Princeton), 1959.

A GENERALIZATION OF A POINCARÉ-BENDIXSON THEOREM TO CLOSED TWO-DIMENSIONAL MANIFOLDS.*

By Arthur J. Schwartz.

Introduction. We shall consider a C^2 action of the real line, T, on a compact, connected, two-dimensional manifold, M, of class C^2. Thus α: $T \times M — M$ is a transformation of class C^2 such that $tm \equiv \alpha(t, m)$ satisfies $(t+s)m = t(sm)$ and $0m = m$. Note that C^2 vector fields on M give rise to such actions (see, fo rexample [5, p. 40]). A set, $\Omega \subset M$, is called α-*minimal* if it is closed, invariant under α $(T\Omega — \Omega)$, non-empty and contains no such proper subset.

The purpose of this paper is to prove that any minimal set is either a $_1$xed point, a homeomorph of S^1, or ll of M and homeomorphic to T^2. This theorem is an extension of work by Poincaré [7], Bendixson [1], and Denjoy [2]. In particular, the analytical methods of Denjoy are modified, extended, and to some degree simplified. A theorem equivalent to this was stated by Haas [3, p. 297]. However, as Peixoto pointed out [6, p. 113], the accompanying proof was incorrect.

The author wishes to acknowledge his indebtedness to Professors R. Sacksteder, E. Lima, and R. Ellis fo rtheir criticism and encouragement during the preparation of this paper.

Theorem. *Let M be a compact, connected, two-dimensional manifold of class C^2. Let α: $T \times M \to M$ be a C^2 action of the reals on M. Let $\Omega \subset M$ be an α-minimal set. Then Ω must be one of the following:* a) *a singleton consisting of a fixed point*; b) *a single, closed orbit homeomorphic to S^1; or* c) *all of M which is homeomorphic to a torus T^2.*

Proof. Let $\Omega \subset M$ be an α-minimal set. Then Ω may be either

(a) $\{\omega\} = T\omega$, i.e., a fixed point;

(b) an orbit, γ, which is homeomorphic to S^1, i.e., a periodic orbit; or

(c) a set which contains neither fixed points nor closed orbits.

* Received April 23, 1963.

Ω cannot be anything else since fixed points and closed orbits, being non-empty, closed, and invariant may not be properly contained in a minimal set. Case (c) may be subdivided as follows:

(c′) Ω has non-empty interior. In this case, since the set of interior points is invariant and Ω is minimal, the set of boundary points must be empty. Thus Ω is open and closed and must be all of M. It follows from a result of Kneser [4, p. 153] that since M contains neither fixed points nor closed orbits it must be homeomorphic to T^2.

(c″) Ω has empty interior and, being closed, is nowhere dense.

Since (a), (b) and (c′) are precisely those cases which are possible according to the conclusion of the theorem, it remains to show that (c″) cannot occur.

Let $i\colon [-1,1] \to M$ be a C^2 imbedding such that

(a) $I = i((-1,1))$ is transverse to every orbit intersecting it;

(b) $i(-1)$ and $i(+1)$ are not in Ω;

(c) $i(0)$ is in Ω.

I may be used as a local cross section; i.e., there exists $\sigma > 0$ such that for $|s| < \sigma$ and $|t| < 1$, $\delta(s,t) = si(t)$ defines a diffeomorphism. We may thus take δ^{-1} as a coordinate system at $i(0)$.

The proof of the theorem turns on the behavior of a function, $f(\tau)$, induced by "the first return of the orbit through $\tau \in I$ to I." We define this function as follows:

Let

$$U = \{x \in I \mid \exists t > 0, tx \in I\}$$

and for each $x \in U$

$$t_x = \min\{t \mid tx \in I, t > 0\},$$

then for $v \in i^{-1}(U) = V \subset (-1,1)$, we define

$$f(v) = i^{-1}(t_{i(v)} i(v)).$$

Note that for $|v - v_0|$ sufficiently small,

$$f(v) = \pi(\delta^{-1}(t_{i(v_0)} i(v))),$$

where $\pi(\xi,\eta) = \eta$, and therefore $f \in C^2(V)$. Since Ω is minimal, for each $\omega \in \Omega$ and real number N, $\omega_N = \mathrm{cl}\{t\omega \mid \omega \in N\} = \Omega$ if "cl" denotes "closure."

In fact, since M is compact, $\bigcap \omega_N$ for $N \geqq 0$ is a closed, invariant subset of Ω. Thus

$$G = i^{-1}(I \cap \Omega) \subset V.$$

It is also true that G is a perfect, nowhere dense set. Letting W be open in $(-1,1)$ and such that $G \subset W \subset \bar{W} \subset V$, we summarize the properties of f:

$$G = (-1,1) - \bigcup_{i=1}^{\infty} (a_i, b_i) \subset W, \tag{1}$$

$$f\colon W \to (-1,1), \tag{2}$$

$$f(G) = G, \tag{3}$$

$$f^k(g) = g \text{ and } g \in G \text{ implies } k = 0, \tag{4}$$

$$0 < L \leqq |f'(w)| \leqq F \text{ for all } w \in W,\ 0 < L < 1 < F, \tag{5}$$

$$|f''(w)| \leqq M \text{ for all } w \in W. \tag{6}$$

$$G \text{ is } B\text{-minimal, where } B(k,g) = f^k(g). \tag{7}$$

We shall prove the theorem by showing that no function satisfying (1)-(7) exists.

Let $\mu = \operatorname{dist}(G, (-1,1) - W)$. We now prove:

LEMMA 1. *There exists a complementary interval (a,b) such that $a < b$; $a, b \in G$; and $f^k((a,b)) \subset W$ for all $k \geqq 0$.*

Proof of Lemma 1. Let $S = \{i \mid b_i - a_i \geqq \mu\}$ and $Y = \{a_i, b_i \mid i \in S\}$. Note that S and Y are finite sets. Thus it follows from (4) that there exists N such that $f^k(a_1) \notin Y$ for $k \geqq N$. Since f preserves endpoints of complementary intervals, $f^N(a_1) = a_i$ or b_i for some i. Then, according to the choice of N, $|f^k(a_i) - f^k(b_i)| < \mu$ for all $k \geqq 0$. Then setting $(a,b) = (a_i, b_i)$, we have an interval with the required properties and the lemma is proved.

Using the symbol D to denote differentiation, $[Dg](x) = g'(x)$, we verify the following:

LEMMA 2. *Let N be an integer and $[p,q] \subset (-1,1)$ such that $f^k([p,q]) \subset W$ for all k satisfying $0 \leqq k \leqq N$. Then*

$$f^{k+1}(p) - f^{k+1}(q) = Df^{k+1}(w) \cdot (p - q) \tag{8}$$

for some $w \in (p,q)$;

$$|Df^{k+1}(u)| \cdot |Df^{k+1}(v)|^{-1} \leqq \exp[ML^{-1} \sum_{j=0}^{k} |f^j(p) - f^j(q)|] \tag{9}$$

for all k satisfying $0 \leqq k \leqq N$ and $u, v \in [p,q]$.

Proof. Formula (8) is derived by applying the mean value theorem to f^{k+1}. Next, since $Df^{k+1}(u) = \prod_{j=0}^{k} Df(f^j(u))$,

$$\begin{aligned}\log[|Df^{k+1}(u)|\cdot|Df^{k+1}(v)|^{-1}] \\ &\leqq \sum_{j=0}^{k} |\log Df(f^j(u)) - \log Df(f^j(v))| \\ &= \sum_{j=0}^{k} |Df(w_j)|^{-1}|D^2f(w_j)|\cdot|f^j(u) - f^j(v)|,\end{aligned}$$

which implies (9).

Next, we note that Lemma 1 implies that

$$\sum_{k=0}^{\infty} Df^k(w_k) \leqq 2(b-a)^{-1} \tag{10}$$

for some $\{w_k\} \subset (a,b)$. Employing (9) and (10) we derive

$$\sigma = \sum_{k=0}^{\infty} Df^k(a) \leqq 2(b-a)^{-1}\exp[2ML^{-1}]. \tag{11}$$

Now set

$$\nu = \mu L[6(\sigma+1)(M+1)]^{-1}. \tag{12}$$

We now assert that for any x in $N(\nu) = \{y \mid |y-a| < \nu\}$ and any natural number k,

$$|f^k(x) - f^k(a)| < \mu \tag{13-k}$$

and

$$|Df^k(x)| < e\,|Df^k(a)|. \tag{14-k}$$

We note first that (13-0) and (14-0) are trivially valid. Suppose (13-k) and (14-k) are valid for $0 \leqq k \leqq N$. Then according to (9), (8), the induction assumption, (11), and (12), respectively, we have

$$\begin{aligned}|Df^{N+1}(x)|\,|Df^{N+1}(a)|^{-1} &\leqq \exp[ML^{-1}\sum_{k=0}^{N}|f^k(x) - f^k(a)|] \\ &\leqq \exp[ML^{-1}\sum_{k=0}^{N} Df^k(u_k)\cdot\nu] \leqq \exp[ML^{-1}e\cdot\sigma\cdot\nu] < e,\end{aligned}$$

since $ML^{-1}e\sigma\nu < e\mu/6 < \mu \leqq 1$. (14 k) and (8) imply (13-k) and the assertion follows by induction.

Now (11) and (14) imply that $\lim Df^k(x) = 0$, $k \to \infty$, uniformly for

x in $N(\nu)$. Moreover, uniform convergence holds for x in $f^p(N(\nu) \cap G)$ since $|Df^k(f^p(x))| \leqq F^{|p|}L^{-|p|}\,|Df^k(x)|$. Since

$$K = \cup\{f^p(N(\nu) \cap G) \mid -\infty < p < \infty\}$$

is a non-empty, open invariant set in G, and G is B-minimal, $K = G$. Since G is compact, we have

$$\lim_{k\to\infty} Df^k(x) = 0 \text{ uniformly in } G. \tag{15}$$

However, our analysis applies equally well to f^{-k} and we must therefore have

$$\lim_{k\to\infty} Df^{-k}(x) = 0 \text{ uniformly in } G. \tag{16}$$

Together, (15) and (16) imply that for some N, $Df^N(x) < \frac{1}{2}$ and $Df^{-N}(f^N(x)) < \frac{1}{2}$ which is absurd. Thus the theorem is proved.

Application to Poincaré-Bendixson theory. The "classical" Poincaré-Bendixson theory is concerned, in large part, with the study of ω-limit sets which are defined as follows:

Definition. Let p be a point in M^2, then the *ω-limit set of p* is

$$\Omega p = \{q \mid \exists t_n \to \infty, t_n p \to q\}.$$

It can be shown in a straightforward manner that $\Omega p = \bigcap_{N\geqq 1} \mathrm{cl}\{tp \mid t \geqq N\}$; whence we may deduce that Ωp is closed, invariant, and non-empty. It follows from an application of Zorn's lemma that Ωp contains at least one minimal set.

Next we introduce the notion of "winding" in a precise fashion.

Definition. A *transversal, $I \subset M^2$,* is a differentiable homeomorph of $[0,1]$ such that no orbit of α intersects I tangentially.

Definition. We say *tp winds toward Ωp* in case for every q in Ωp, there is a transversal, Iq, containing q in its interior such that successive intersections of $\{tp \mid t \geqq 0\}$ with Iq tend monotonically toward q.

We may now prove the following "Poincaré-Bendixson type" theorem:

COROLLARY. Let *M^2 be an orientable manifold. If M^2 is not a minimal set under a C^2 action, α, and $\Omega p \subset M^2$ contains no fixed points, then $\Omega p \approx S^1$ and tp winds toward Ωp.*

Proof. ΩP contains at least one minimal set, Σ. From the preceding theorem it follows that $\Sigma \approx S^1$. We may construct, around Σ, a neighborhood,

K, consisting of a union of transversals, which is homeomorphic to a cylinder, $(-1,1)\times S^1$. It follows from the continuity of α, and the fact that $\Sigma\subset\Omega p$, that there exist $0<t_1<t_2$, $q\in\Omega p$, and Iq a transversal through q such that t_2p is between t_1p and q on Iq. Thenceforth tp is "trapped" between the closed curve formed by $[t_1,t_2]p$ and the portion of Iq between t_1p and t_2p and Σ. Thus we find the next intersection of tp with Iq between t_2p and q. Repetition of this argument yields the conclusion of the corollary.

PRINCETON UNIVERSITY.

REFERENCES.

[1] I. Bendixson, "Sur les courbes définies par des équations différentielles," *Acta Mathematica*, vol. 24 (1901), pp. 1-88.

[2] A. Denjoy, "Sur les courbes définies par les équations différentielles à la surface du tore," *Journal de Mathématiques Pures et Appliquées*, vol. 11 (1932), pp. 333-375.

[3] F. Haas, "Poincaré-Bendixson type theorems for two dimensional manifolds different from the torus," *Annals of Mathematics*, vol. 59 (1954), pp. 292-299.

[4] H. Kneser, "Regulare Kurvenscharen auf den Ringflächen," *Mathematische Annalen*, vol. 91 (1924), pp. 135-154.

[5] S. Lefschetz, *Differential equations: geometric theory*, Interscience, New York, 1957.

[6] M. M. Peixoto, "Structural stability on two dimensional manifolds," *Topology*, vol. 1 (1962), pp. 101-120.

[7] H. Poincaré, "Sur les courbes définies par les équations différentielles," *Comptes Rendus de l'Académie des Sciences*, Paris, vol. 90 (1880), pp. 673-675.

SIDE APPROXIMATION, MISSING AN ARC.*

By David S. Gillman.[1]

I. Introduction. Much topological study has been devoted to embeddings of 2-spheres in Euclidean 3-dimensional space E^3. By a *2-sphere,* we mean any homeomorphic image of the boundary of a 3-cell. The 2-sphere S is said to be *tamely embedded in* E^3 (or simply *tame*) if there is a homeomorphism h of E^3 onto itself such that $h(S)$ is a polyhedron (the sum of a locally finite collection of tetrahedra, triangles, segments, and points). If S is not tame, then S is called *wild.* Tameness and wildness may be defined in this manner for any closed subset of E^n. In 1926, Alexander exhibited the first example of a wild 2-sphere in E^3 [1]; many examples have appeared subsequently.

Indeed, these examples illustrate that a 2-sphere embedded in E^3 may have many surprising characteristics besides that of wildness—it may have several points at which it cannot be pierced by a tame arc [11]—it may well have some subarcs which are wildly embedded in E^3, as does the 2-sphere of Alexander—it may have all of its subarcs tamely embedded, although the 2-sphere itself is wildly embedded [5].

On the other hand, recent results of Bing severely limit the amount of pathology that may occur in a 2-sphere S in E^3. It is now known that S must contain many tame arcs [6]; it also must contain many points at which it can be pierced by a tame arc [7]. Bing established these properties by repeated use of his Side Approximation Theorem [8], which we call Theorem 1. It say the following:

Theorem 1. *If S is a 2-sphere, U is a component of $E^3 - S$, and $\epsilon > 0$, then there is a polyhedral 2-sphere S', a finite collection $D_1, D_2, \cdots, D_k$ of mutually exclusive disks lying in S', each of diameter less than or equal to ϵ, a finite collection $E_1, E_2, \cdots, E_m$ of mutually exclusive disks lying in S, each of diameter less than or equal to ϵ, such that*

1. *there is a homeomorphism of S onto S' that moves no point by more than ϵ (henceforth abbreviated $H(S, S') \leqq \epsilon$).*

*Received July 5, 1963.

[1] This paper formed part of the author's dissertation at the University of Wisconsin under the direction of Professor R. H. Bing. This research was supported by a National Science Foundation fellowship.

2. $S' - \sum_{i=1}^{k} D_i \subset U$.

3. $S \cdot S' \subset \sum_{i=1}^{m} E_i$.

It would seem, then, that the subdisks $E_1, E_2, \cdots, E_m$ are, in some sense, the "bad" points of S. This provides the motivation for the main theorem of this paper, a characterization of tame arcs in E^3 (Theorem 10). This theorem states that a necessary and sufficient condition that an arc A, lying in a 2-sphere S in E^3, be tame is that Theorem 1 is still true for S with the additional condition:

4. $A \cdot \sum_{i=1}^{m} E_i = \emptyset$.

What are some of the possible applications of this theorem? Consider the following questions raised by Bing:

Question 1. If A is a tame arc lying in a 2-sphere S in E^3, can S be pierced by a tame arc at each point of A? [7]

Question 2. Suppose Y is the set of points of S at which S cannot be pierced by a tame arc. Is Y an 0-dimensional F_σ set? [7]
In the course of this paper, both questions will be answered in the affirmative. We now outline the arrangement of the material to follow.

In Section II, only the necessity of the condition characterizing tame arcs is established, this being done in Theorem 2. The main purpose of the third section is to develop, via Theorem 2, a characterization of the piercing points of a 2-manifold in E^3—a 2-manifold can be pierced by a tame arc at p if and only if there is a tame arc lying in the 2-manifold and containing p. Note that one direction of this result answers Question 1 in the affirmative.

Sections IV, V, and VI are devoted primarily to a proof of the sufficiency of the condition characterizing tame arcs. Once this result is established, it, together with the piercing point characterization of Section III, is then used to yield an affirmative answer to Question 2. The best previously known result in the direction of Question 2 was given in [7], showing that the set Y lies in a 0-dimensional G_δ subset of S.

II. Side approximation, missing a tame arc. We make the following definitions:

The *interior* of a 2-sphere S, abbreviated as $Int\, S$, is the bounded com-

plementary domain of S. This term will also be used in another sense, namely, the *interior* of a disk D, abbreviated as $Int\, D$, consists of the points of D not lying on the rim of D.

The distance between two sets V and W will be denoted by $dist(V, W)$; the diameter of a set X will be denoted by $diam(X)$.

An ϵ *set* is a set of diameter less than or equal to ϵ.

A set X is *locally polyhedral* at a point p of E^3 is there is a neighborhood N of P such that $\bar{N} \cdot X$ is a polyhedron in E^3.

A set X is *locally tame* at a point p of E^3 if there is a neighborhood N of p and a homoeomorphism h of $\bar{N}$ into E^3 such that $h(\bar{N} \cdot X)$ is a polyhedron in E^3.

THEOREM 2. *If S is a 2-sphere in E^3, A is a tame arc lying in S, U is a component of $E^3 - S$, and $\epsilon > 0$, then there is a polyhedral 2-sphere S', a finite collection $D_1, D_2, \cdots, D_k$ of mutually exclusive ϵ disks lying in S', a finite collection $E_1, E_2, \cdots, E_m$ of mutually exclusive ϵ disks lying in S, such that*

1. $H(S, S') \leqq \epsilon$.
2. $S - \sum_{i=1}^{k} D_i \subset U$.
3. $S \cdot S' \subset \sum_{i=1}^{m} E_i$.
4. $A \cdot \sum_{i=1}^{m} E_i = \emptyset$.

Proof. Observe that Conditions 1, 2, and 3 of Theorem 2 are those of Theorem 1. The main tool employed in this proof is a strong form of Theorem 1, namely, the Side Approximation Theorem on Open Subsets, established in [8]. For convenience we assume that the component U of $E^3 - S$ is Int S.

Given $\epsilon > 0$, we assume that S has diameter greater than 2ϵ and that there is a point p of Int S whose distance from S is more than 2ϵ. There is a $\delta > 0$ such that if f is a homeomorphism of S moving points less than δ, then any δ subset of $f(S)$ will lie on a disk in $f(S)$ of diameter less than ϵ.

We now invoke the Side Approximation Theorem for the open set $S - A$, and for the positive number $\delta/2$ (Theorem 18 of [8]). This tells us that there is a homeomorphism g defined on S such that

a. g is fixed on A.
b. g moves points less than $\delta/2$.
c. $g(S)$ is locally polyhedral, mod A.

d. $g(S-A)$ lies, except for an infinite collection $D_1^*, D_2^*, \cdots$, of mutually exclusive $\delta/2$ disks, in $\operatorname{Int} S$.
e. $S-A$ lies, except for an infinite collection $E_1^*, E_2^*, \cdots$, of mutually exclusive $\delta/2$ disks, in $\operatorname{Ext} g(S)$.
f. Both collections of disks are locally finite, mod A.

Since the 2-sphere $g(S)$ is locally polyhedral, mod A, where A is a tame arc, it follows that $g(S)$ is tame [10]. Thus, by [2], there exists a homeomorphism k of E^3 onto itself taking $g(S)$ onto the unit sphere R.

Let α be a homeomorphism of E^3 onto itself which is the identity except in a small neighborhood of R, and which contracts R slightly, i.e., $\alpha(R) \subset \operatorname{Int} R$. We choose α to move points a sufficiently small amout so that the homeomorphism $[k^{-1}\alpha k]$ moves points less than $\delta/2$. The 2-sphere $S' = [k^{-1}\alpha k]g(S)$ will now be shown to satisfy Conditions 1-4 of Theorem 2. Of course, S' may not be polyhedral; this will be attended to later.

1. By its very definition, it follows that S' is homeomorphically within $\delta/2 + \delta/2 = \delta$ of S. Since $\delta < \epsilon$, Condition 1 follows.

The proof that S' satisfies Condition 2 will be deferred until after the proofs of Conditions 3 and 4.

3. $S' = [k^{-1}\alpha k]g(S)$ is contained in $\operatorname{Int} g(S)$. Thus, Property e above implies that

$$S - A - \sum_{i=1}^{\infty} E_i^* \subset \operatorname{Ext} S'.$$

Furthermore, since $A \subset g(S)$, we may strengthen this to

$$S - \sum_{i=1}^{\infty} E_i^* \subset \operatorname{Ext} S'.$$

Since the disks $E_1^*, E_2^*, \cdots$, converge to A, there is an integer N such that

$$S - \sum_{i=1}^{N} E_i^* \subset \operatorname{Ext} S'.$$

Thus, the disks $E_1^*, E_2^*, \cdots, E_N^*$ satisfy Condition 3.

4. The disks $E_1^*, E_2^*, \cdots, E_N^*$ were selected to lie in $S-A$, so Condition 4 is immediate.

2. Any component of $S \cdot S'$ is a δ set, since it lies in a disk E_i^*, which is a δ set. Furthermore, since S' is homeomorphically within δ of S,

it follows that every component of $S \cdot S'$ lies in a disk in S' of diameter less than ϵ.

Let $D_1, D_2, \cdots, D_k$ be a finite collection of mutually exclusive ϵ disks in S' such that

$$S \cdot S' \subset \sum_{i=1}^{k} D_i.$$

That there exists such a collection is shown in Theorem 9 of [8]. Furthermore,

$$S' - \sum_{i=1}^{k} D_i \subset \operatorname{Int} S.$$

The last relationship may be seen as follows: Note that p lies in $\operatorname{Int} S'$, since S' is homeomorphically within δ of S. This is a consequence of Theorem VI 13 of [13]. Select a point q lying in $\operatorname{Ext} S'$ such that the distance from q to $S + S'$ is greater than ϵ. Thus, by Theorem 14 of [8], there is an arc pq with end points p and q such that

$$pq \cdot \left[\sum_{i=1}^{k} D_i + \sum_{i=1}^{N} E_i^* \right] = \emptyset.$$

The first point of $[S' - \sum D_i] + [S - \sum E_i^*]$ on pq in the order from p to q is a point of S' since

$$S - \sum_{i=1}^{N} E_i^* \subset \operatorname{Ext} S'.$$

Moreover, this point must lie in $\operatorname{Int} S$, since p was chosen to lie in $\operatorname{Int} S$. The desired relationship follows, completing the proof of Condition 2.

Thus, the 2-sphere S' satisfies Conditions 1-4; however, although S' is tame, it is not necessarily polyhedral, as we would like. This may be remedied by use of [3] or [4], choosing a polyhedral 2-sphere sufficiently close to S' so that Conditions 1-4 remain true. This completes the proof of Theorem 2.

A *Sierpinski Curve* X is the set obtained by removing from a 2-sphere the interiors of a sequence of mutually exclusive disks, dense on the 2-sphere, and with diameters converging to zero. The boundary points of these disks are called *accessible points* of X. The remaining points of X are called *inacccessible.*

It is known that any two Sierpinski Curves are homeomorphic.

The main result established in [6] is the following.

THEOREM 3. *For each 2-sphere S in E^3 and each positive number ϵ there is a Sierpinski Curve X in S such that*

1) *X lies in a tame 2-sphere, and*
2) *each component of $S - X$ is of diameter less than ϵ.*

This result is established in [6] by repeated use of Theorem 1. By making repeated use of Theorem 2 in place of Theorem 1, the methods of [6] establish the following stronger result.

THEOREM 4. *If A is a tame arc lying in a 2-sphere S in E^3, then for each positive number ϵ, there is a Sierpinski Curve X in S such that*

1) *X lies in a tame 2-sphere*
2) *each component of $S - X$ is of diameter less than ϵ*
3) *A lies in X; in fact, A lies in the set of inaccessible points of X.*

This argument will establish the same result for subsets of 2-spheres other than tame arcs; in fact, the argument is valid for any closed set C for which it is known that a 2-sphere which is locally polyhedral, mod C, is tame.

Question. Can C be any set which lies in a tame 2-sphere and which does not have a degenerate component?

COROLLARY 4.1. *If A is a tame arc lying in a 2-sphere S in E^3, then there is a tame simple closed curve J such that $A \subset J \subset S$.*

Proof. By Theorem 4, there is a Sierpinski Curve X which lies in a tame 2-sphere, which lies in S, and which contains A in its inaccessible points. Let G be the upper semicontinuous decomposition of S whose nondegenerate elements are the disks of $S - X$. Since A lies in the inaccessible points of X, A does not intersect any of the countable collection of nondegenerate elements of G. Furthermore, by Moore's theorem [15], the decomposition space is a 2-sphere. It is easy to find a simple closed curve J in the decomposition space which contains the image of A in the decomposition space, and such that J does not intersect any nondegenerate element in the decomposition space. The preimage of J will contain A, will lie in S, and will be tame, since it lies entirely in X, which in turn lies in a tame 2-sphere in E^3.

III. Piercing surfaces with tame arcs. An arc is said to *pierce* a 2-sphere S in E^3 if the arc intersects S at only one point, and the ends of the arc lie in different components of $E^3 - S$. Other definitions of piercing are sometimes used. If D is a disk in E^3 and apb is an arc that intersects D only

at the point p of $\operatorname{Int} D$, then apb *pierces* D at p if and only if no small neighborhood of p contains an arc from ap to bp that misses p. Another alternative definition would be that apb *pierces* D if apb lies in a simple closed curve J such that each neighborhood of p contains another simple closed curve linking J. These three definitions are all equivalent, and will be used interchangeably. For a complete discussion of linking, see [3].

In [7] Bing used Theorem 3 to show that every 2-sphere S in E^3 can be pierced by a tame arc. In fact, the argument showed even more, namely, if A is an arc lying in a tame Sierpinski Curve $X \subset S$ such that A lies in the inaccessible points of X, then there is a disjoint family of tame arcs piercing S at each point of A. Using this fact and Theorem 4, therefore, we are able to conclude the following.

THEOREM 5. *If A is a tame arc lying in a 2-sphere S in E^3, then S can be pierced by a tame arc at every point of A. In fact, there is a disjoint family of tame arcs which pierce S simultaneously at each point of A.*

This theorem answers in the affirmative Question 1 of the Introduction.

Theorem 5 may be generalized to disks in a manner similar to that used in [7]. We define a *cellular set* in E^3 to b a set which is the common intersection of a decreasing sequence of 3-cells. If C is a cellular set lying in the interior of a disk D in E^3, then the argument of Theorem 2.1 of [7] establishes the existence of a subdisk D' of D such that

1) C is contained in the interior of D'

2) D' lies in a 2-sphere in E^3.

Since any tame arc is a cellular set, the problem of dealing with a tame arc lying in the interior of a disk in E^3 reduces to that of a tame arc lying in a 2-sphere. This establishes the following.

COROLLARY 5.1. *If A is a tame arc lying in any 2-manifold M in E^3, then M can be pierced by a tame arc at every point of A. In fact, there is a disjoint family of tame arcs which pierce M simultaneously at each point of A.*

A *spanning arc* of a disk is an arc whose end points belong to the boundary of the disk and whose non-end points belong to the interior of the disk.

LEMMA 6.1. *If A is an arc in a 2-sphere S in E^3, p is a point of A such that A is locally tame,* mod p, *and S can be pierced by a tame arc at p, then A is tame.*

Proof. By Theorem 9 of [4], we assume that A is locally polyhedral, mod p. By Theorem 7 of [3], we further assume that S is locally polyhedral, mod A. Thus, the only points of S that may not be locally tame points of S are points of A. Let q be any point of $A-p$; let D be a small disk of S, containing q on its interior, such that p does not lie in D, and such that $D\cdot A$ is a spanning arc of D. By Theorem 1 of [12], D is tame. Hence S is locally tame at q.

The only point, therefore, where S is not known to be locally tame is p. But S can be pierced at p by a tame arc, so Theorem 1 of [14] implies that S is tame. Since A lies in a tame 2-sphere, it follows that A itself is a tame arc.

LEMMA 6.2. *If A_1 and A_2 are tame arcs intersecting in a common end point b, and if A_1+A_2 lies in a 2-sphere S in E^3, then A_1+A_2 is tame.*

Proof. By Theorem 5, it follows that S can be pierced by a tame arc at b. Thus, the arc A_1+A_2 satisfies the hypothesis of Lemma 6.1, so A_1+A_2 is tame.

That all of the hypotheses of Lemma 6.1 and Lemma 1.2 are actually necessary may be seen from examples of wild arcs given in [11].

We now are able to characterize the piercing points of a 2-sphere.

THEOREM 6. *The 2-sphere S can be pierced at a point p by a tame arc B if and only if there is a tame arc A lying in S and containing p.*

Proof. One direction follows immediately from Theorem 6. The other direction is handled thus:

Given the point p and tame arc B piercing S at p, we may construct a sequence $\{A_i\}$ of tame arc in S such that $A_i\cdot A_{i+1}\neq\emptyset$ for all i, and $\lim_{i\to\infty} A_i=p$. This sequence is built by repeated use of Theorem 3. A sequence of subarcs $A'_i\subset A_i$ is now selected so that $A'_i\cdot A'_j=\emptyset$ if $|i-j|\geqq 2$, and $A'_i\cdot A'_{i+1}$ is a common endpoint of both arcs. The arc $A=p+\sum A'_i$ is locally tame at points of the form $A'_i\cdot A'_{i+1}$ by Lemma 6.2. Thus, A is tame, by Lemma 6.1.

We note that although the arc A was constructed to contain the point p only as an end point, this is really no restriction in view of Corollary 4.1.

COROLLARY 6.1. *If A is an arc in a 2-manifold M in E^3, p is a point of A such that A is locally tame,* mod p, *and M can be pierced by a tame arc at p, then A is tame.*

COROLLARY 6.2. *If A_1 and A_2 are tame arcs intersecting in a common end point b, and if A_1+A_2 lies in a 2-manifold in E^3, then A_1+A_2 is tame.*

Corollary 6.3. *A 2-manifold M can be pierced at a point p by a tame arc B if and only if there is a tame arc A lying in M and containing p.*

These corollaries are established exactly as was Corollary 5.1, that is, by placing a section of the 2-manifold M on a 2-sphere.

IV. A simple closed curve which pierces many disks. The primary goal of the next three sections is to complete the characterization of tame arcs, one direction of which has already been established in Theorem 2. The problem may be stated thusly: If X is a closed subset of a 2-sphere S in E^3, with X and S possessing the following property, which we indicate by $(*, X, S)$ for convenience, then what can be said about the set X?

Property $(*, X, S)$. If U is a component of $E^3 - S$, and ϵ is a positive number, then there is a polyhedral 2-sphere S', a finite collection $D_1, D_2, \cdots, D_k$ of mutually exclusive ϵ disks lying in S', a finite collection $E_1, E_2, \cdots, E_m$ of mutually exclusive ϵ disks lying in S, such that

1. $H(S, S') < \epsilon$.
2. $S' - \sum_{i=1}^{k} D_i \subset U$.
3. $S \cdot S' \subset \sum_{i=1}^{m} E_i$.
4. $X \cdot \sum_{i=1}^{m} E_i = \emptyset$.

In particular, if $X = A$ is an arc, it is true that $(*, A, S)$ implies that A is tame? The natural approach to the problem requires the following.

Conjecture. $(*, X, S)$ is a local property of the set X and 2-sphere S. That is, if S_1 is another 2-sphere in E^3 such that there exists a disk D with $X \subset \operatorname{Int} D \subset D \subset S \cdot S_1$, and if $(*, X, S)$ is satisfied, then $(*, X, S_1)$ is satisfied.

Being unable to establish this conjecture, we are forced to prove that $(*, A, S)$ implies that the arc A is tame via a considerably more devious method.

Theorem 7. *If A is an arc lying in a 2-sphere S in E^3 and satisfying $(*, A, S)$, then there exists a simple closed curve J in S with $A \subset J$ such that $(*, J, S)$ is satisfied.*

Proof. Let J_1 be any simple closed curve containing A and lying in S. Using $(*, A, S)$, one may approximate S from $U = \operatorname{Int} S$ within ϵ_1 so that

$A \cdot \sum_{i=1}^{m} E_i = \emptyset$. A homeomorphism h_1 is now defined on J_1, fixed on A, and taking $J_1 - A$ entirely into $S - \sum_{i=1}^{m} E_i$. The approximation process is now repeated, this time from $\operatorname{Ext} S$ within ϵ_2, and a homeomorphism h_2 is defined on $J_2 = h_1(J_1)$, with h_2 fixed on A, and moving $J_2 - A$ off of the second set of ϵ subdisks of S. If this procedure is repeated again and again, alternating the approximations to be first from $\operatorname{Int} S$ and then from $\operatorname{Ext} S$, and if the ϵ_i are selected sufficiently small, then $J = \lim J_i$ will satisfy the conclusion of Theorem 10.

The ϵ_i must be chosen, firstly so that

$$\{\cdots h_3 h_2 h_1\} : J_1 \to J$$

is actually a homeomorphism, by means of Theorem 7 of [7]. Secondly, we require that ϵ_2 is less than one fourth the distance from J_2 to the set of ϵ_1 disks; we require that ϵ_3 is less than one fourth the distance from J_3 to the set of ϵ_2 disks, as well as $\epsilon_3 < \epsilon_2/2$. In general, we require that ϵ_i is less than one fourth the distance from J_i to the set of ϵ_{i-1} disks, as well as $\epsilon_i < \min(\epsilon_{i-1}/2, \epsilon_{i-2}/4, \cdots, \epsilon_2/2^{i-2})$. This insures that J is disjoint from all ϵ_i disks, $i = 1, 2, \cdots$, and completes the proof.

COROLLARY 7.1. *If p is a point lying in a 2-sphere S in E^3 and satisfying $(*, p, S)$, then there exists a simple closed curve J in S with p contained in J such that $(*, J, S)$ is satisfied.*

The proof is identical to that of Theorem 7.

LEMMA 8.1. *If J is a simple closed curve lying in a 2-sphere S in E^3, and p is a point of J at which S can be pierced by a tame arc, then there is a tame arc A in S which "crosses" J at p; that is, $A \cdot J = p$, and $A - p$ has two components which lie in different complementary domains of $S - J$.*

Proof. By the technique used to prove Theorem 6, in each complementary domain of $S - J$, a tame arc may be built having p as an end point. The union of these two tame arcs is tame, by Lemma 6.2.

THEOREM 8. *If J is a simple closed curve lying in a 2-sphere in E^3 and satisfying $(*, J, S)$, then there is a disk E such that J pierces E.*

Proof. The proof is broken into six steps.

1) If $U = \operatorname{Int} S$ and $\epsilon > 0$, then $(*, J, S)$ assures the existence of a polyhedral 2-sphere S' satisfying four conditions. We observe that for ϵ sufficiently small, the third condition may be replaced by a stronger one:

$$3' \cdot S - \sum_{i=1}^{m} E_i \subset \operatorname{Ext} S'.$$

The argument showing this fact is given in the proof of Condition 3 of Theorem 2, as well as in the proof of Theorem 4 of [6], and will not be repeated here. As a consequence of Condition 3′ and Condition 4, we have that $J \subset \operatorname{Ext} S'$. The analogous result holds for $U = \operatorname{Ext} S$, of course.

2) Given $\epsilon > 0$, there is a 2-sphere Q satisfying the four conditions of $(*, J, S)$ for $U = \operatorname{Int} S$ and ϵ, and another 2-sphere R satisfying the four conditions of $(*, J, S)$ for $U = \operatorname{Ext} S$ and ϵ. We now show that we may assume that $Q \cdot R = \emptyset$.

We may suppose that the diameter of J is greater than ϵ. There exists a positive number δ_1 such that any δ_1 subset of S lies in an ϵ disk of S. There exists a positive number δ_2 such that any δ_2 subset of a 2-sphere lying homeomorphically within δ_2 of S lies in a $\delta_1/4$ disk of that 2-sphere.

There exists a polyhedral 2-sphere Q', a collection of δ_2 disks $D'_1, D'_2, \cdots, D'_k$ in Q', such that

1. $H(S, Q') < \delta_2$.
2. $Q' - \sum_{i=1}^{k} D'_i \subset \operatorname{Int} S$.

Letting δ_3 be the minimum of δ_2 and $\operatorname{dist}(Q' - \sum D'_i, S)$, there exists a polyhedral 2-sphere R satisfying the four conditions of $(*, J, S)$ for $U = \operatorname{Ext} S$ and δ_3. Thus, $R \cdot [Q' - \sum D'_i] = \emptyset$. The 2-sphere Q' is now adjusted slightly on the subdisks D'_i to form a new 2-sphere Q which will be entirely disjoint from R, and which still satisfies the conditions of $(*, J, S)$ for $U = \operatorname{Int} S$ and ϵ. This adjustment is performed as follows.

Without loss of generality, we suppose that Q' and R are in general position. Let K be an "innermost" simple closed curve in R of all the simple closed curves of $[\sum_{i=1}^{k} D'_i] \cdot R$, i.e., K bounds a small subdisk Z of R with $[\operatorname{Int} Z] \cdot Q' = \emptyset$. Since K lies on some D'_j, it follows that $\operatorname{diam}(K) < \delta_2$, so $\operatorname{diam}(Z) < \delta_1/4$. We replace the subdisk bounded by K in D'_j by the subdisk Z of R, then push this adjusted subdisk slightly to one side of R, thus eliminating at least one simple closed curve of $[\sum_{i=1}^{k} D'_i] \cdot R$. This process is continued until all intersection is eliminated. The new 2-sphere obtained in this manner is called Q and the newly obtained subdisks of Q derived from the disks $\sum D'_i$ are called $\sum D_i$. Obviously, $Q \cdot R = \emptyset$. All that remains to be seen is that Q satisfies Conditions 1-4 of $(*, J, S)$ for $U = \operatorname{Int} S$ and ϵ.

It is easy to verify that Q is homeomorphically within $\delta_1/2$ of Q'. Thus, Q is homeomorphically within $\delta_1/2 + \delta_2$ of S, so $H(S, Q) < \epsilon$.

Each disk D_i has diameter less than $\operatorname{diam}(D'_i) + \delta_1/2$, which is less

than δ_1. The relationship $Q' - \sum_{i=1}^{k} D'_i \subset \operatorname{Int} S$ immediately shows, moreover, that $Q - \sum_{i=1}^{k} D_i \subset \operatorname{Int} S$.

The ϵ disks of S satisfying Conditions 3 and 4 still need to be constructed. Since each component of $Q \cdot S$ lies on some D_j, it has diameter less than δ_1, and, therefore, lies in an ϵ disk of S. We take the upper semicontinuous decomposition G of S whose nondegenerate elements are components of $Q \cdot S$, then consider the big 2-sphere of the resulting cactoid. The arc A, when viewed in the decomposition space, misses the image of these nondegenerate elements, and must lie in the big 2-sphere, since $\operatorname{diam}(A) > \epsilon$. Thus when small disks are selected covering the image of these nondegenerate elements, these disks may be selected to miss A. The preimage of these disks in S are the desired disks $E_1, E_2, \cdots, E_m$. For more details, see Theorem 9 of [8].

3) We now build a first approximation of the disk E. Let p be any point of J at which S can be pierced by a tame arc; it is known that there are many such points [7]. By Lemma 8.1, there is a tame arc A crossing J at the point p. Now Theorem 4 may be used to find a Sierpinski Curve X such that X lies in a tame 2-sphere and A lies in the inaccessible points of X. By moving this tame 2-sphere with a space homeomorphism, we may assume that X lies in the xy-plane in E^3, that A is the straight line interval $\{(x,y,z)/x=0,\ -1 \leqq y \leqq \pm 1,\ z=0\}$, and that $p=(0,0,0)$.

The set X is obtained from S by removing the interiors of a sequence of disjoint disks $B_1, B_2, \cdots$. We select a disk E_1 in the yz-place such that $\operatorname{Int} E_1$ contains $(0,0,0)$, A is a spanning arc of E_1, $[\operatorname{Bd} E_1] \cdot [\sum B_i] = \emptyset$, and $\operatorname{Bd} E_1$ links J; this is shown schematically in the figure.

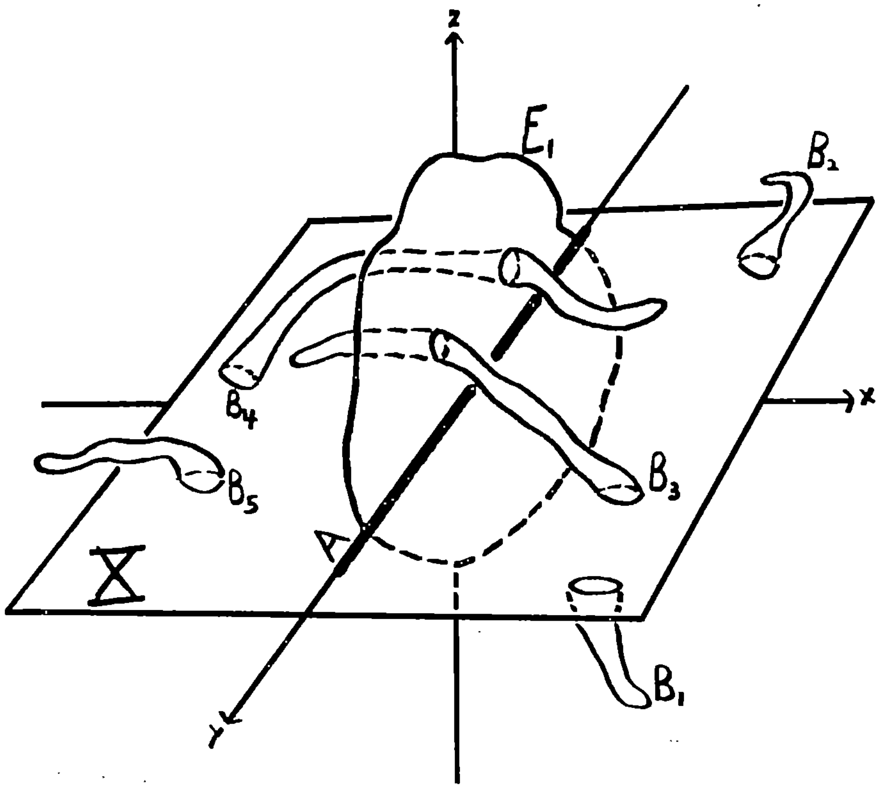
The disk E_1 is a first approximation to the desired disk E. Of course, E_1 may not yet satisfy the conclusion of Theorem 8, because J may intersect E_1 many times on feelers like B_3 and B_4 in the figure. By means of $(*, J, S)$, we will adjust E_1 to eliminate such intersection, thus completing the proof. In the figure, it is clear that E_1 can be deformed around B_3 and B_4; however, if B_3 and B_4 had "horns" which hooked together as in [1], and if J ran through these horns, then the situation would be very bad. Intuitively speaking, this is exactly what $(*, J, S)$ prevents from happening.

4) Let i_1 be the smallest integer such that $B_{i_1} \cdot J \cdot E_1 \neq \emptyset$. In this step, we alter E_1 to remove such intersection. We define

$$\alpha = \operatorname{dist}(E_1 \cdot B_{i_1}, [S - B_{i_1}] \cdot E_1 + \operatorname{Bd} E_1)$$
$$\mathcal{O} = \{x \mid x \text{ lies in } E_1 \text{ and } \operatorname{dist}(x, E_1 \cdot B_{i_1}) < \alpha/3\}$$
$$\partial\mathcal{O} = \{x \mid x \text{ lies in } E_1 \text{ and } \operatorname{dist}(x, E_1 \cdot B_{i_1}) = \alpha/3\}.$$

The alteration of E_1 will occur only in the set $\mathcal{O}$.

There exists a polyhedral 2-sphere Q and a polyhedral 2-sphere R satisfying the conditions of $(*, J, S)$ for $U = \operatorname{Int} S$ and for $U = \operatorname{Ext} S$, respectively, and for the positive number ϵ_1. The size of ϵ_1 will be decided on later. By Step 1, we may assume $J \subset \operatorname{Ext} Q$ and $J \subset \operatorname{Int} R$. By Step 2, we may assume $Q \cdot R = \emptyset$. Call q the homeomorphism taking S onto Q and moving points no more than ϵ_1; call r the analogous homeomorphism taking S onto R.



Since $\partial\mathcal{O} \cdot S = \emptyset$, by choosing ϵ_1 sufficiently small, we insure that all points of $\partial\mathcal{O}$ lie either in $\operatorname{Int} Q \cdot \operatorname{Int} R$ or in $\operatorname{Ext} Q \cdot \operatorname{Ext} R$. This follows from Theorem VI 13 of [13]. Assuming that Q, R, and E_1 are in general position, we label the simple closed curves of $[Q + R] \cdot \mathcal{O}$ as $K_1, K_2, \cdots, K_t$. By choosing ϵ_1 sufficiently small, we may require that

$$\sum_{i=1}^{t} K_i = q(B_{t_1}) \cdot E_1 + r(B_{t_1}) \cdot E_1.$$

Select an "innermost" K_j in $q(B_{t_1})$ which we will relabel as K_1 for convenience. Replace the disk that it bounds in E_1 with the disk that it

bounds in $q(B_{i_1})$, then deform this new disk slightly to one side of $q(B_{i_1})$, thus eliminating the intersection. This process is continued, relabeling a second, now "innermost" simple closed curve as K_2, replacing a disk, and deforming to one side, as before. When all the simple closed curves K_i which lie in $q(B_{i_1})$ have been removed, the same process is carried out on $r(B_{i_1})$. We call the disk constructed from E_1 in this manner E_2.

Since Q and R are both disjoint from J, we can guarantee that $J \cdot E_2 \subset J \cdot E_1$. Furthermore, the accomplishment of this adjustment is the following.

$$B_{i_1} \cdot J \cdot E_2 = \emptyset.$$

To see this, suppose that y is a point of $B_{i_1} \cdot J \cdot E_2$. Since $J \subset \operatorname{Ext} Q \cdot \operatorname{Int} R$, any ray from y to ∞ in general position will intersect $Q + R$ an odd number of times. We now construct a ray contradicting this fact. The point y lies in $\mathcal{O}$, yet does not lie in the interior of any simple closed curve K_j, or it would have been eliminated in the formation of E_2 from E_1. It follows that an arc may be drawn in E_1 from y to a point of $\partial \mathcal{O}$, so that this arc is disjoint from $Q + R$. Since each point of $\partial \mathcal{O}$ lies either in $\operatorname{Ext} Q \cdot \operatorname{Ext} R$ or in $\operatorname{Int} Q \cdot \operatorname{Int} R$, we have the desired contradiction.

We note that $(0,0,0)$ is not moved in forming E_2 from E_1, because all alteration took place in the set $\mathcal{O}$; $\mathcal{O}$ does not intersect the arc A, while A spans E_1 and contains the point $(0,0,0)$.

5) We select i_2 to be the smallest integer such that $B_{i_2} \cdot J \cdot E_2 \neq \emptyset$, then approximate S within ϵ_2, and adjust E_2 to remove this intersection, exactly as in the previous step, forming E_3. The process is continued and the set $E = \lim_{i \to \infty} E_i$ will now be shown to satisfy the conclusion of Theorem 8.

6) By observing that $\lim_{j \to \infty} B_{i_j} = (0,0,0)$, and requiring that $\lim_{i \to \infty} \epsilon_i = 0$, we insure that E is actually a disk. Furthermore, J intersects E only at $(0,0,0)$, and J links $\operatorname{Bd} E = \operatorname{Bd} E_1$; thus J pierces E.

COROLLARY 8.1. *If J is a simple closed curve lying in a 2-sphere S in E^3 and satisfying $(*, J, S)$, then there is a dense subset of J at each point of which J pierces a disk.*

Proof. In the proof of Theorem 8, the point p where J pierced the disk E was selected with only one restriction: It was a point of J at which S could be pierced by a tame arc (Step 3). Since such points form a dense subset of J [7], the result follows.

V. Surrounding J with thin tori. A *torus* is any set homeomorphic to the cartesian product of two simple closed curves. A *solid torus* is any

set homeomorphic to the cartesian product of a 2-cell and a simple closed curve. If T is a torus in E^3, it is not always true that $T + \operatorname{Int} T$ is a solid torus, for $T + \operatorname{Int} T$ may be a cube with a knotted hole instead. In the statement of Theorem 9, the concept of solid torus will not be used at all; instead, the weaker concept of a 2-dimensional torus plus its interior will be used exclusively, for convenience in the proof. The theorem is also true when stated using solid tori in place of tori plus their interiors; this follows trivially once Theorem 10 is established.

THEOREM 9. *If J is a simple closed curve lying in a 2-sphere S and satisfying $(*, J, S)$, then J is the intersection of a decreasing sequence of tori plus their interiors.*

Proof. It suffices to show that, given $\epsilon > 0$, there exists a torus T in E^3 such that $J \subset \operatorname{Int} T$, and T lies in an ϵ neighborhood of J (the set of all points in E^3 whose distance from J is less than ϵ). Roughly, the idea of the proof is this: Select a small annulus on S which contains J. Use $(*, J, S)$ to deform this annulus into $\operatorname{Ext} S$ and into $\operatorname{Int} S$. Lastly, join the edges of these two deformed annuli with two more annuli, thus forming a torus. We now give details.

We may select simple closed curves J_1 and J_2 in S such that J_1 and J_2 are in opposite complementary domains of $S - J$, J_1 and J_2 are tame, and the annulus V which lies in S and has $J_1 + J_2$ as its boundary will lie in an ϵ neighborhood of S. This is easily done via Theorem 3. Without loss of generality we assume that J_1 and J_2 are polygonal.

A simple closed curve K may be found in $E^3 - V$ which links J_1 and J_2, i.e., which links the annulus V. Let R be a ray starting at any point j of J, such that $R \cdot S = j$. We assume that R is locally polyhedral mod j. J_1 and J_2 are now blown up a bit into thin disjoint tori T_1 and T_2 such that

(a) J_i circles $\operatorname{Int} T_i$ once longitudinally, $i = 1, 2$.

(b) $J \subset \operatorname{Ext} T_i$, $i = 1, 2$.

(c) $T_1 + T_2 + V$ has diameter less than ϵ.

(d) $K \cdot [T_i + \operatorname{Int} T_i] = \emptyset$, $i = 1, 2$.

(e) $[T_i + \operatorname{Int} T_i]$ intersects only one complementary domain of $S - J$, $i = 1, 2$.

(f) $T_i \cdot R = \emptyset$, $i = 1, 2$.

Inside the torus T_i, even a thinner torus T'_i is selected so that J'_i circles $\operatorname{Int} T'_i$ once longitudinally, and thin enough so that if D is a disk of V with $\operatorname{Bd} D \subset T'_i + \operatorname{Int} T'_i$, then $D \subset T_i + \operatorname{Int} T_i$.

The annulus V will now be deformed a small amount to either side of S. By $(*, J, S)$ there exist homeomorphisms q and r of S into E^3 such that

(1) $J \subset \operatorname{Ext} q(S)$ and $J \subset \operatorname{Int} r(S)$.

(2) $q(S)$ and $r(S)$ are disjoint.

That there are such homeomorphisms follows exactly from the proofs of Steps 1 and 2 of Theorem 8. We place six more restrictions on $q(S)$ and $r(S)$, all of which may be obtained simply by requiring that points are moved a sufficiently small amount by q and r.

(3) $q(J_i)$ and $r(J_i)$ both circle $\operatorname{Int} T'_i$ once longitudinally, $i = 1, 2$.

(4) $T_1 + T_2 + q(V) + r(V)$ lies in an ϵ neighborhood of J.

(5) If D is a disk of $q(V)$ or $r(V)$, and $\operatorname{Bd} D \subset T'_i$, then $D \subset T_i$, $i = 1, 2$.

(6) K links the annuli $q(V)$ and $r(V)$.

(7) $[T'_i + \operatorname{Int} T'_i]$ intersects only one complementary domain of $q(S) - q(J)$, $i = 1, 2$. The same requirement is made for r.

(8) Let B be a small annulus contained in $V - T_1 - T_2$ which contains J. We require that q and r move points less than $\operatorname{dist}(B, T_1 + T_2)$ and $\operatorname{dist}(R, S - B)$.

Without loss of generality, we assume that $q(V)$, $r(V)$, T'_1, T'_2, and R are in general position. Thus, $q(V) \cdot [T'_1 + T'_2]$ is a finite collection of simple closed curves. If L is one of these simple closed curves, then L is either longitudinal on the annulus $q(V)$ or bounds a disk on $q(V)$. Also, L cannot be meridianal on T'_i, since it would link $q(J_i)$; thus L is either longitudinal on T'_i or bounds a disk on T'_i, by Theorem 1 of [9]. We now show that L is longitudinal on $q(V)$ if and only if it is longitudinal on T'_i.

If L is longitudinal on T'_i, then it cannot be shrunk to a point in $E^3 - K$. Also, if L is longitudinal on $q(V)$, it cannot be shrunk to a point in $E^3 - K$. But if L bounds a disk on either $q(V)$ or on T'_i, then it can be shrunk to a point in $E^3 - K$. The desired correspondence follows.

We now remove all longitudinal simple closed curves of $q(V) \cdot [T'_1 + T'_2]$ from $q(V)$, dividing $q(V)$ into components. Call C' the component of $q(V)$ containing $q(J)$; C' will have two boundary simple closed curves which, because of Restriction (7) above, cannot both lie on the same torus T'_i. We select C'' analogously in $r(V)$. Restriction (8) implies that

$$R \cdot q(S) \subset R \cdot q(B) \subset R \cdot C'$$

and

$$R \cdot r(S) \subset R \cdot r(B) \subset R \cdot C''$$

Since R intersects $q(S)+r(S)$ an odd number of times, R must intersect $C'+C''$ an odd number of times. Selecting an annulus U_1 on T'_1 joining C' to C'', an annulus U_2 on T'_2 joining C' to C'', the set $C'+C''+U_1+U_2$ forms a singular torus intersecting R an odd number of times. We now change this into a real torus.

An "innermost" simple closed curve M of the singular set of $C'+C''+U_1+U_2$ is selected on U_1. M bounds a disk in either C' or C''; this disk is replaced by the disk bounded by M in U_1, then moved slightly to the side, thus reducing the singularity. This process is continued until a real torus T is formed. By Restriction (4) above, T will lie in an ϵ neighborhood of J. The ray R will intersect T an odd number of times; for, each replacement in yielding T from the singular torus is made by interchanging two disks, both of which lie in $T_i+\operatorname{Int} T_i$, $i=1,2$, by Restriction (5). Thus, R does not intersect either of these two disks, and $R\cdot T$ has an odd number of points. This implies that $J\subset \operatorname{Int} T$.

VI. Completing the tame arc characterization. We may now prove the main result of this paper.

THEOREM 10. *An arc A lying in a 2-sphere S in E^3 is tame if and only if $(*,A,S)$ is satisfied.*

Proof. One direction has already been proved (Theorem 2). We now establish the other direction.

By Theorem 7, A is contained in a simple closed curve J such that $(*,J,S)$ is satisfied. Corollary 8.1 and Theorem 9 tells us that J satisfies the hypothesis of Theorem 2 of [12]. Thus J is tame.

THEOREM 11. *If S is a 2-sphere and Y is the set of points of S at which S cannot be pierced by a tame arc, then Y is a 0-dimensional F_σ set (and, therefore, of the 1st category).*

Proof. That Y is 0-dimensional follows from [7]. We define Y_i to be the set of all points p in S such that $(*,p,S)$ cannot be satisfied for $U=\operatorname{Int} S$ and $\epsilon=1/i$. Similarly, Z_i is defined to be the set of all points p in S such that $(*,p,S)$ cannot be satisfied for $U=\operatorname{Ext} S$ and $\epsilon=1/i$. For each positive integer i, Y_i and Z_i are closed sets. Furthermore, each Y_i and Z_i is nowhere dense in S, since every tame arc of S lies in $S-Y_i-Z_i$, by Theorem 2. This shows that $\sum Y_i+\sum Z_i$ is a 0-dimensional F_σ set, and, therefore, of the 1st category in S.

We complete the proof by showing that $Y=\sum Y_i+\sum Z_i$. If p lies in

$S - Y$, then p lies in a tame arc A in S, by Theorem 6. But A does not intersect any Y_i or Z_i, by Theorem 2. Thus, p lies in $S - (\sum Y_i + \sum Z_i)$. Conversely, if p lies in $S - (\sum Y_i + \sum Z_i)$, then $(*, p, S)$ is satisfied. By Corollary 7.1 and Theorem 10, p lies in a tame arc A in S. By Theorem 6, p does not lie in Y.

UNIVERSITY OF WISCONSIN AND
THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] J. W. Alexander, "An example of a simply connected surface bounding a region which is not simply connected," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 10 (1942), pp. 8-10.

[2] ———, "On the subdivision of 3-space by a polyhedron," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 10 (1924), pp. 6-8.

[3] R. H. Bing, "Approximating surfaces with polyhedral ones," *Annals of Mathematics*, vol. 65 (1957), pp. 456-483.

[4] ———, "Locally tame sets are tame," *Annals of Mathematics*, vol. 59 (1954), pp. 145-158.

[5] ———, "A wild surface, each of whose arcs is tame," *Duke Mathematical Journal*, vol. 28 (1961), pp. 1-15.

[6] ———, "Each disk in E^3 contains a tame arc," *American Journal of Mathematics*, vol. 84 (1962), pp. 583-590.

[7] ———, "Each disk in E^3 *is pierced by a tame arc*," *American Journal of Mathematics*, vol. 84 (1962), pp. 591-599.

[8] ———, "Approximating surfaces from the side," *Annals of Mathematics*, vol. 77 (1963), pp. 145-192.

[9] ———, "Point-like decomposition of E^3," *Fundamenta Mathematicae*, vol. 50 (1962) pp. 431-453.

[10] P. H. Doyle and J. G. Hocking, "Some results on tame disks and spheres in E^3," *Proceedings of the American Mathematical Society*, vol. 11 (1960), pp. 832-836.

[11] R. H. Fox and E. Artin, "Some wild cells and spheres in three-dimensional spaces," *Annals of Mathematics*, vol. 49 (1948), pp. 979-990.

[12] D. S. Gillman, "Sequentially 1-*ULC* tori," Transactions of the American Mathematical Society, to appear.

[13] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton, 1948.

[14] E. E. Moise, "Affine structures in 3-manifolds, VII, Disks which are pierced by intervals," *Annals of Mathematics*, vol. 58 (1953), pp. 403-408.

[15] R. L. Moore, "Concerning upper semicontinuous collections of continua, *Transactions of the American Mathematical Society*, vol. 27 (1925), pp. 416-428.

THE STRUCTURE OF DISTAL FLOWS.*

By H. Furstenberg.[1]

Introduction. We shall be dealing exclusively with flows on a compact metric space. By a flow on X one usually understands a (continuous) representation of the additive group of reals as a group of continuous transformations $\{\tau_t, -\infty < t < \infty\}$ of the space X. Before stating the problem we shall study, it will be best to summarize what is known for a rather special kind of flow. We say that a flow is *equicontinuous* if the transformations τ_t form an equicontinuous family. In the case of equicontinuous flows, two assertions can be made, and these completely determine the structure of such flows. First, if there is on X an equicontinuous flow, then X decomposes into a disjoint union of subsets, each of which is invariant under the τ_t, and on each of which the flow is *minimal*. We say a flow is minimal on a space if no closed proper subset of the space is carried into itself by all the transformations of the flow. The analysis of equicontinuous flows is thereby reduced to that of minimal equicontinuous flows. The second assertion concerns the latter. Namely, if there is a minimal equicontinuous flow on X, then X must be homeomorphic to a compact abelian group, and the transformations of the flow correspond to multiplication in the group by elements of a one-parameter subgroup. Conversely, starting with a compact abelian group and a one-parameter subgroup, there is determined a flow on the group which is equicontinuous. The theory of equicontinuous flows is thereby reduced to the theory of compact abelian groups.

It should perhaps be pointed out that these facts have a bearing on the theory of almost periodic functions (in the sense of Bohr) on the real line. If $f(t)$ is an almost periodic function, then the uniform closure of the translates of f is a compact space X, and the operations of translation in X determine a flow which is clearly equicontinuous in the uniform metric on X. This flow can be seen to be minimal and so X is a compact abelian group. The real line is imbedded as a dense subgroup of X and one may show that $f(t)$ extends to a continuous function on X. The theory of compact abelian groups then implies, among other things, that f is a uniform limit of linear com-

* Received April 22, 1963.

[1] This research was sponsored in part by the AirForce Office of Scientific Research.

binations of characters on X; hence $f(t)$ is a uniform limit of linear combinations of exponentials on $(-\infty, \infty)$.

We turn now to the (informal) definition of a distal flow. Let $d(x,y)$ denote the metric on the space X, and let $\{\tau_t\}$ be a 1-parameter group of transformations of X. We say the corresponding flow is *distal* if for any pair of distinct points $x, y \in X$, the distances $d(\tau_t x, \tau_t y)$ are bounded away from 0 for $-\infty < t < \infty$. It is not hard to see that this definition is independent of the metric used. Now if a flow is equicontinuous, we can always introduce an equivalent metric $d'(x,y)$, for which $d'(\tau_t x, \tau_t y)$ is independent of t; so an equicontinuous flow is always distal. For some time it was believed that the converse might be true, but now there are a number of examples showing that this is not the case. We will presently examine some of these examples.

Once it is recognized that the class of distal flows is larger than that of equicontinuous ones, we may ask whether the results mentioned above have an analogue for distal flows. The first mentioned property (which is by no means trivially true) is in fact known to be valid and a general distal flow decomposes into minimal ones. This a consequence of the work of Ellis ([4]). The problem we shall study is to find an analogue of the second assertion; that is, to exhibit, in a sense, the most general minimal distal flow.

One class of examples of distal flows that are not equicontinuous is given by the "nilflows" or flows on nilmanifolds ([1, 2]). Let N be a nilpotent Lie group and Γ a discrete subgroup such that N/Γ is compact. The group N acts in a natural way on the homogeneous space N/Γ, and if $\{g(t)\}$ is a one-parameter subgroup of N, there is determined a flow on N/Γ. This flow is known to be distal, and in general it will not be equicontinuous.

Another class of examples, but of discrete distal flows (with the group of integers replacing the reals) is given in [6]. A particular example is the flow on the 2-torus (ξ, η), $|\xi| = |\eta| = 1$ generated by

$$\tau(\xi, \eta) = (e^{i\alpha}\xi, \psi(\xi)\eta) \tag{1}$$

where ψ is a continuous function with $|\psi(\xi)| = 1$. Let us show that this is distal.

The map $(\xi, \eta) \to \xi$ gives a fibering of the torus over the circle into a bundle of circles. The transformation τ permutes the fibres among themselves. Moreover, each fibre (i. e. for fixed ξ) is transformed isometrically, so that pairs of points on the same fibre remain the same distance apart under the flow. Different fibres also remain the same distance apart under the flow, so two points on distinct fibres can never get closer than the distance between these fibres. This make it clear that the flow is distal.

On the other hand, the flow generated by (1) is generally not equicontinuous. For example, take $\psi(\xi) \equiv \xi$. Then

$$\tau^n(\xi, \eta) = (e^{in\alpha}\xi, e^{in(n-1)\alpha/2}\xi^n\eta). \tag{2}$$

In particular the second component of $\tau^n(e^{\pi i/n}\xi, \eta)$ is the negative of that of $\tau^n(\xi, \eta)$. Hence these two points are bounded away from one another although $(e^{\pi i/n}\xi, \eta) \to (\xi, \eta)$. Hence the family $\{\tau^n\}$ is not equicontinuous.

This second class of examples can be enlarged considerably. Suppose we are already given a distal flow in a space X. Let Y be an m-sphere bundle over X with the orthogonal group in $m+1$ variables as the structure group. Let a flow on Y be given by having τ_t take the sphere above $y \in Y$ into the sphere above $\tau_t y$, the transformations of the individual spheres being rotations. More generally, instead of m-spheres, we may take homogeneous spaces of arbitrary compact groups. The reasoning used before shows that such a flow on Y will again be distal. Now the same procedure may be repeated finitely, or even infinitely, often, and starting with an equicontinuous flow, we may build up a large family of distal flows in this manner.

Indeed, the flows on nilmanifolds also arise in this way. For example, let N be a 2-step nilpotent Lie group, i.e., N/N_1 is commutative, where N_1 is the centralizer of N. Let Γ be a discrete subgroup of N with N/Γ compact. Then $N/N_1 \cdot \Gamma$ is a torus and there is a natural map $N/\Gamma \to N/N_1 \cdot \Gamma$. N/Γ is then a fibre bundle over a torus and the fibres are $\approx N_1 \cdot \Gamma/\Gamma \approx N_1/N_1 \cap \Gamma$ and since N_1 is commutative, the latter is a compact abelian group. It is not hard to verify that the flow on N/Γ arises from a equicontinuous flow on the torus $N/N_1 \cdot \Gamma$ in the manner described above.

Our main result, Theorem 2.4, is suggested by these observations. Namely, we shall show that by approximately this procedure, one may build up the most general minimal distal flow. This result will be stated precisely in § 2.

At this point, the author wishes to acknowledge his indebtedness to Frank Hahn who pointed out certain similarities between the distal flows and equicontinuous flows and who posed the problem of analyzing the structure of the former.

1. Preliminaries. We shall consider flows for groups more general than the real line or the integers. The arguments are virtually the same for an arbitrary locally compact group as for the reals and we have one application for the general result (§ 12) which is uninteresting in the case of 1-parameter flows. The locally compact parameter group will be denoted by T and will be fixed throughout the discussion.

Definition 1. A *flow* (X, T) on the compact metric space X is given by a continuous map of $T \times X \to X$, denoted $(\tau, x) \to \tau x$, satisfying $\tau_1(\tau_2 x) = (\tau_1\tau_2)x$, where $\tau_1, \tau_2 \in T$, $x \in X$, and $ex = x$ where e is the identity element of T.

Definition 1.2. Let (X, T) and (Y, T) be two flows and let π be a map of X onto Y. We say that (Y, T) is a *subflow* of (X, T) ((X, T) is an *extension* of (Y, T)) relative to π, if for every $\tau \in T$, $x \in X$, $\pi(\tau x) = \tau(\pi(x))$. We also write in this case $(Y, T) = \pi(X, T)$.

Let (Y, T) be a subflow of (X, T) relative to π. If $y \in Y$, we shall sometimes denote the inverse image $\pi^{-1}(y)$ by X_y. The condition that (Y, T) be a subflow of (X, T) may be expressed as $\tau X_y \subset X_{\tau y}$, $\tau \in T$, $y \in Y$. The subsets X_y will be referred to as the *fibres* of X (with respect to Y, or over Y).

Definition 1.3. Two flows (X, T) and (Y, T) are *isomorphic* if (Y, T) is a subflow of (X, T) relative to π, and the map $\pi\colon X \to Y$ is a homeomorphism. If π is not a homeomorphism, we shall say that (Y, T) is a *proper* subflow of (X, T), or that (X, T) is a *non-trivial* extension of (Y, T).

Definition 1.4. The flow (X_0, T) is *trivial* if X_0 consists of a single point, and the transformations of T are all identity transformations.

The trivial flow is a subflow of every flow.

Let (X, T) be a flow and $\pi_1(X, T) = (Y_1, T)$, $\pi_2(X, T) = (Y_2, T)$ two subflows. Suppose, that there is a map $\pi'\colon Y_1 \to Y_2$ such that $\pi_2 = \pi' \circ \pi_1$. We claim that if such a map π' exists, it is uniquely determined by this condition. For $\pi_1\colon X \to Y_1$ is onto, so that π' is determined by the requirement $\pi_2 = \pi' \circ \pi_1$. Now in most of our discussion a flow (X, T) will be given such that all other flows that occur are subflows of this one. If we speak of a subflow (Y, T) of (X, T) it is to be understood that there is a fixed map $\pi_Y\colon X \to Y$ relative to which (Y, T) is a subflow of (X, T). If for two such subflows, (Y_1, T) and (Y_2, T) we speak of (Y_2, T) as a subflow of (Y_1, T), it will then be with respect to the unique map π' of Y_1 onto Y_2 satisfying $\pi_{Y_2} = \pi' \circ \pi_{Y_1}$. If $y_1 \in Y_1$, then $y_1 = \pi_{Y_1}(x)$, $x \in X$, and $\tau y_1 = \pi_{Y_1}(\tau x)$. Hence

$$\pi'(\tau y_1) = \pi' \circ \pi_{Y_1}(\tau x) = \pi_{Y_2}(\tau x) = \tau\pi_{Y_2}(x) = \tau\pi' \circ \pi_{Y_1}(x) = \tau\pi'(y_1)$$

and π' actually does determine (Y_2, T) as a subflow of (Y_1, T). As a result there is determined a partial ordering among the subflows of (X, T):

$$\pi_2(X, T) \prec \pi_1(X, T) \text{ if } \pi_2 = \pi' \circ \pi_1$$

for some π'.

Definition 1.5. The flow (X,T) is *minimal* if no closed proper subset of X is invariant under the action of T.

Equivalently, (X,T) is minimal if Tx is dense in X for every $x \in X$. If (X,T) is any flow, there will exist (by the compactness of X) at least one minimal closed subset $X' \subset X$ invariant under T. There is then a flow (X',T) on X' and this is clearly minimal. We speak of (X',T) as a *minimal component* of (X,T), and of X' as a *minimal set* in X. The minimal sets of X do not always exhaust the space X. For example, let X be the one point compactification of the real line, $X = R \cup \{\infty\}$, let $T = R$ with $\tau_t(s) = s + t$, $\tau_t(\infty) = \infty$. Then $\{\infty\}$ is the only minimal set in X. An important property of the distal flows that we shall study is that they are "semi-simple," in the sense that X is the union of its minimal sets (which are, of course, non-overlapping).

Definition 1.6. Let $d(x_1, x_2)$ denote a metric on X. The flow (X,T) is *distal* if whenever x_1, x_2 are distinct points of X, $d(\tau x_1, \tau x_2)$ is bounded away from 0 as τ ranges over T.

Equivalently, (X,T) is distal unless for a pair of distinct points $x_1, x_2 \in X$ there is a sequence τ_n in T and a point $z \in X$ such that $\tau_n x_1 \to z$, and $\tau_n x_2 \to z$. From this it follows that the above definition is independent of the metric $d(\cdot,\cdot)$.

2. Isometric extensions. Let M be a homogeneous compact metric space. By this we mean a compact metric space such that for any two points $x, y \in M$, there is an isometry of M taking x into y. The isometries of M form a compact group H, and M may be identified with a coset space H/H_0, where H_0 is the subgroup of H leaving a given point of M fixed. Conversely, if H is a compact group satisfying the first axiom of countability and H_0 is a closed subgroup, then H/H_0 may be given a metric which is invariant under H, so that $M = H/H_0$ is a homogeneous compact metric space.

Definition 2.1. Let X and Y be compact metric spaces, π a map of X onto Y, and M a homogeneous compact metric space. We say that X is an *M-bundle over* Y if there is a real valued function $\rho(x_1, x_2)$ defined for all pairs x_1, x_2 in the same fibre of X (i.e. whenever $\pi(x_1) = \pi(x_2)$) and such that

(a) $\rho(x_1, x_2)$ is continuous as a function on the subset of $X \times X$ defined by the condition $\pi(x_1) = \pi(x_2)$.

(b) For each $y \in Y$, $\rho(x_1, x_2)$ defines a metric on the fibre X_y under which X_y is isometric to M.

For example, let X be a fibre bundle over Y with fibre M and with

structure group, the group of isometries of M, and let π be the map of $X \to Y$. If U is a small enough open subset of Y, then $\pi^{-1}(U) \approx U \times M$, and we can define $\rho_U(x_1, x_2)$ for $\pi(x_1) = \pi(x_2) \in U$ as the metric in M. Let $\{U_i\}$ be a covering of Y by such neighborhoods. On $\pi^{-1}(U_i) \cap \pi^{-1}(U_j)$ the functions $\rho_{U_i}(x_1, x_2)$ and $\rho_{U_j}(x_1, x_2)$ agree since the structure group acts by isometries. It is then evident that X is an M-bundle over Y. It is not clear whether the notion of an M-bundle is actually more general than that of a fibre bundle of the kind just described.

Definition 2.2. Let X be an M-bundle over Y with $\pi\colon X \to Y$, let (X, T) be a flow and $(Y, T) = \pi(X, T)$ a subflow. We say (X, T) is an *isometric extension* of (Y, T) if the function $\rho(x_1, x_2)$ of Def. 2.1 satisfies $\rho(\tau x_1, \tau x_2) = \rho(x_1, x_2)$ for all $\tau \in T$, and x_1, x_2 in the same fibre of X over Y.

Thus each fibre of X is transformed isometrically by the transformations of T. The following proposition was proved informally in the Introduction.

PROPOSITION 2.1. *An isometric extension of a distal flow is distal.*

Proof. Let $\pi\colon X \to Y$ where (X, T) is an isometric extension of (Y, T). If (X, T) is not distal, then for some pair $x_1 \neq x_2$, and sequence $\{\tau_n\}$ in T, $\tau_n x_1 \to z$, $\tau_n x_2 \to z$. Applying π to these limits, we have $\tau_n \pi(x_1) \to \pi(z)$, $\tau_n \pi(x_2) \to \pi(z)$. But since (Y, T) is distal, this can only happen if $\pi(x_1) = \pi(x_2)$. Then $\rho(\tau_n x_1, \tau_n x_2)$ is defined and equals $\rho(x_1, x_2)$, But $\rho(\tau_n x_1, \tau_n x_2) \to \rho(z, z) = 0$ by the continuity of ρ. This gives $\rho(x_1, x_2) = 0$ and since ρ defines a metric on each fibre of X, $x_1 = x_2$. This is a contradiction and so (X, T) is distal.

Definition 2.3. Let (X, T) be a flow and let $\Sigma = \{(X_\alpha, T)\}$ be a collection of subflows of (X, T), $(X_\alpha, T) = \pi_\alpha(X, T)$. We say that (X, T) is a *limit* of the flows in Σ in case for every pair of distinct points $x_1, x_2 \in X$ there is a subflow (X_α, T) such that $\pi_\alpha(x_1) \neq \pi_\alpha(x_2)$.

For example let X denote the infinite dimensional torus: $X = \{(\zeta_1, \zeta_2, \cdots), |\zeta_i| = 1\}$, let X_n denote the n-dimensional torus and let $\pi_n\colon X \to X_n$ be defined by $\pi_n(\zeta_1, \zeta_2, \cdots) = (\zeta_1, \zeta_2, \cdots, \zeta_n)$. If T denotes the integers and (X, T) is generated by

$$\tau(\zeta_1, \zeta_2, \cdots, \zeta_n, \cdots) = (\phi_1(\zeta_1), \phi_2(\zeta_1, \zeta_2), \cdots, \phi_n(\zeta_1, \cdots, \zeta_n), \cdots), \tag{2.1}$$

then (X, T) is a limit of $(X_n, T) = \pi_n(X, T)$.

PROPOSITION 2.2. *A limit of distal flows is distal.*

Proof. Let (X,T) be a limit of a collection $\Sigma = \{(X_\alpha, T)\}$ of subflows. We suppose that each (X_α, T) is distal. Suppose (X,T) were not distal. Then for a pair x_1, x_2 of distinct points of X we would have $\tau_n x_1 \to z$, $\tau_n x_2 \to z$. Now for some (X_α, T), $\pi_\alpha(x_1) \neq \pi_\alpha(x_2)$. On the other hand $\tau_n \pi_\alpha(x_1) \to \pi_\alpha(z)$, $\tau_n \pi_\alpha(x_2) \to \pi_\alpha(z)$, which is impossible if (X_α, T) is distal.

Definition 2.4. Let (X,T) be a flow and (Y,T) a subflow. Suppose there is an ordinal η such that to each ordinal $\xi \leqq \eta$ there is associated a subflow (X_ξ, T) of (X,T) such that the following are satisfied:

(a) $(X_0, T) = (Y,T)$ and $(X_\eta, T) = (X,T)$ where 0 denotes the least ordinal.
(b) If $\xi < \xi'$, $(X_\xi, T) \prec (X_{\xi'}, T)$ (see the remarks preceding Def. 1.5).
(c) For each $\xi < \eta$, $(X_{\xi+1}, T)$ is an isometric extension of (X_ξ, T).
(d) If ξ is a limit ordinal $\leqq \eta$, then (X_ξ, T) is a limit of $\{(X_{\xi'}, T), \xi' < \xi\}$.

We then say that (X,T) is a *quasi-isometric extension* of (Y,T).

Definition 2.5. A quasi-isometric extension of a trivial flow is called a *quasi-isometric flow* (q. i. flow).

THEOREM 2.3. *Every quasi-isometric flow is distal.*

Proof. If (X,T) is a q. i. flow, then for some ordinal η, $(X,T) = (X_\eta, T)$, where $\{(X_\xi, T), \xi \leqq \eta\}$ is a well-ordered system of subflows of (X,T) satisfying (b), (c), and (d) above with (X_0, T) the trivial flow. Let S denote the set of ordinals $\xi \leqq \eta$ for which (X_ξ, T) is distal. $0 \in S$ and by Proposition 2.1, if $\xi \in S$ then $\xi + 1 \in S$. Also if ξ is a limit ordinal $\leqq \eta$ and $\xi' \in S$ for all $\xi' < \xi$, then $\xi \in S$ by Proposition 2.2. By transfinite induction we conclude that every $\xi \leqq \eta$ is in S and so (X,T) is distal.

We may now give the statement of our main result. It will be proved in § 10.

THEOREM 2.4. *Every minimal distal flow is quasi-isometric.*

Returning to Def. 2.5, we may say that a q. i. flow is one obtained from the trivial flow by a succession of isometric extensions. In our definition we have allowed for an infinite succession. As we shall prove in § 13, this is necessary if we wish to account for all distal flows. As an example consider the flow on the infinite dimensional torus generated by

$$\tau_1(\zeta_1, \zeta_2, \cdots, \zeta_n, \cdots) = (e^{i\alpha}\zeta_1, \zeta_1\zeta_2, \zeta_2\zeta_3, \cdots, \zeta_{n-1}\zeta_n, \cdots)$$

where α is an irrational multiple of π. Defining the subflows (X_n, T) as before,

we see that (X_{n+1}, T) is an isometric extension of (X_n, T). Also the given flow (X, T) is a limit of $\{(X_n, T)\}$. The flow then corresponds to the ordinal ω. We shall prove in § 13 that no smaller ordinal will suffice.

3. The Ellis group of a distal flow. The basic tool used in studying distal flows is the introduction of a certain compactification of the group T. We shall call this the Ellis group after Ellis who first introduced this compactification ([4]). For the reader's convenience we shall reproduce that portion of [4] which is relevant to our disclussion.

Let X be a compact metric space with metric $d(\cdot,\cdot)$. By X^X is meant the space of all transformations of $X \to X$, continuous or not. X^X is a compact Hausdorff topological space with respect to the topology defined as the weakest one rendering the functions $\hat{x}(f) = f(x)$, $f \in X^X$, $x \in X$, continuous. A subbasis for this topology is given by the sets of the form

$$U_\epsilon(x,y) = \{f \in X^x \colon d(x, f(y)) < \epsilon\}$$

where $x, y \in X$, $\epsilon > 0$. Note that for a fixed $f_0 \in X^X$, the map $f \to ff_0$ is continuous on X^X, the inverse image under this map of $U_\epsilon(x,y)$ being $U_\epsilon(x, f_0(y))$. Also the map $f \to f(x_0)$ is by definition continuous, for $x_0 \in X$. On the other hand, the map $f \to f_0 f$ is not in general continuous. However, if f_0 is a continuous transformation of X, then one can show that $f \to f_0 f$ is also continuous.

Now suppose that (X, T) is a flow. We assume (primarily for convenience of notation) that T is effective, so that T may be imbedded in X^X. We let G be the closure of T in X^X; then G is compact.

THEOREM 3.1 (Ellis). *If (X, T) is distal, then G is a group.*

Proof. Let us first show that G is a semigroup. Let $\tau \in T$. The map $f \to \tau^{-1}$ is continuous on X^X (since τ^{-1} is a continuous transformation on X), so $\tau^{-1}G$ is compact. Since $T \subset \tau^{-1}G$, it follows that $G \subset \tau^{-1}G$ or $\tau G \subset G$. Now let $g \in G$. Then $\tau g \in G$ for all $\tau \in T$, so $Tg \subset G$. Since $f \to fg$ is continuous on X^X, it follows that $\overline{Tg} = \bar{T}g = Gg$, so $Gg \subset G$, and we see that G is a semigroup. To prove the theorem we still have to show that every element of G has an inverse in G. Now it is quite easy to see that each $g \in G$ must be 1-1 on X, if (X, T) is distal. For, $d(f(x), f(y))$ is a continuous function on X^X, and if for some $g \in \bar{T}$, $d(g(x), g(y)) = 0$, then the infimum of $d(\tau x, \tau y)$, for $\tau \in T$, must be 0, and so $x = y$ if (X, T) is distal. We will use this observation to prove that each $g \in G$ has a two-sided inverse in G. Let $g \in G$ and form $\Gamma = Gg$. $\Gamma \subset G$, and Γ is a compact semigroup. We shall show that Γ contains an idempotent, i. e., an f with $f^2 = f$. Let A be a minimal

compact subset of Γ satisfying $A \cdot A \subset A$. A exists by Zorn's lemma. Now if $f \in A$, then Af is again compact and $Af \cdot Af \subset Af$. So $Af = A$. Therefore there exists $u \in A$ with $uf = f$. The set of all these u form a compact set $U \subset A$, and $U^2 \subset U$. So again $U = A$ and $f \in U$ whence $f^2 = f$. We claim that f is the identity transformation of X. For $f(f(x)) = f(x)$ for $x \in X$. Since $f \in G$, it is 1-1 and therefore $f(x) = x$ for all $x \in X$ and f is the identity. Moreover we may write $f = g'g$ where $g' \in G$. We claim that g' is a two-sided inverse of g. Since $g'g$ is the identity, $g'gg' = g'$ and $g'gg'(x) = g'(x)$. By the univalence of g', we have $gg'(x) = x$. Hence $g'g = gg' =$ identity. This proves the theorem.

Definition 3.1. The group G will be called the *Ellis group* of the distal flow (X, T). The relative topology it has as a subset of X^X will be termed the E-topology.

One application of Theorem 3.1 is the proof of the "semi-simplicity" of distal flows.

THEOREM 3.2. *If (X, T) is distal then X is the disjoint union of the minimal sets in X of the flow.*

Proof. That the minimal sets of X are disjoint is clear. Now let G be the Ellis group of the flow. We claim that the minimal sets are exactly the orbits Gx for $x \in X$. The sets Gx are invariant since $T \subset G$, so we need only show that no proper closed subset of Gx is invariant under the flow. Now since the map $g \to gy$, $y \in X$, is continuous from G to X, it follows that $\bar{T}y = \overline{Ty}$ or $Gy = \overline{Ty}$. So if y belongs to an invariant closed subset of Gx, then $\overline{Ty} = Gy$ belongs to it as well. But if $y \in Gx$ then $Gy = Gx$ and this proves that Gx is minimal. Clearly the sets Gx exhaust X.

COROLLARY. *If (X, T) is minimal distal, then the Ellis group G is transitive on X.*

The Ellis group is not a topological group in the usual sense. The product $g_1 g_2$ is not even continuous as a function of the single variable g_2, although it is a continuous function of g_1. The operation $g \to g^{-1}$ is therefore also not continuous, in general. It may be shown, in fact, that the Ellis group will be a topological group if and only if the flow (X, T) is equicontinuous.

Also, when the distal flow (X, T) is not equicontinuous, the Ellis group should be thought of as being quite large, in some sense. For example, let T be the integers and X the 2-torus, and let the flow be generated by

$$\tau_1(\zeta_1, \zeta_2) = (e^{i\alpha}\zeta_1, \psi(\zeta_1)\zeta_2) \tag{3.1}$$

where α is an irrational multiple of π. Then the group G consists of all transformations

$$g(\zeta_1,\zeta_2)=(e^{i\beta}\zeta_1,\phi(\zeta_1)\zeta_2) \tag{3.2}$$

where β is real and ϕ is any function of ζ_1 with $|\phi(\zeta_1)|=1$, and subject only to the restriction that

$$\frac{\phi(e^{i\alpha}\zeta)}{\phi(\zeta)}=\frac{\psi(e^{i\beta}\zeta)}{\psi(\zeta)}. \tag{3.3}$$

(3.3) would be a severe restriction, if ϕ were required to be continuous. As it is, ϕ may be multiplied by any function on the circle group of unit modulus which is constant on cosets of the subgroup $\{e^{in\alpha}\}$.

Let (X,T) be a distal flow and let $(Y,T)=\pi(X,T)$ be a subflow.

LEMMA 3.1. *Let G be the Ellis group of (X,T) and let $g\in G$. If, for $x_1,x_2\in X$, $\pi x_1=\pi x_2$, then $\pi g x_1=\pi g x_2$.*

Proof. Since T is dense in G, the set $\{(\tau x_1,\tau x_2),\tau\in T\}$ is dense in $\{(g'x_1,g'x_2),g'\in G\}$. Thus there exists a sequence $\{\tau_n\}$ in T with $\tau_n x_1\to gx_1$, $\tau_n x_2\to gx_2$. Then

$$\begin{aligned}\pi g x_1=\lim \pi\tau_n x_1=\lim \tau_n\pi x_1=\lim \tau_n\pi x_2=\lim \pi\tau_n x_2\\ =\pi g x_2.\end{aligned}$$

As a consequence of this lemma we may map G into Y^Y: set $\bar g(\pi x)=\pi(gx)$ for $g\in G$. This is unambiguous since $\pi x_1=\pi x_2$ implies $\pi(gx_1)=\pi(gx_2)$. The map $g\to\bar g$ is continuous from $G\to Y^Y$ inasmuch as $\{g:\bar g(\pi(x))\in U\}=\{g: gx\in\pi^{-1}U\}$, and this is open if U is open in X. Now T is imbedded in both X^X and Y^Y (possibly after identifying identical transformations) and G is the closure of T in X^X. We conclude that $\bar G$ is the closure of T in Y^Y. Since G is again a group this implies that (Y,T) is distal. For $d(\bar g y_1,\bar g y_2)$ attains its minimum in $\bar G$, where d now denotes a metric in Y. If this minimum were 0, then $\bar g y_1=\bar g y_2$ and $y_1=y_2$. So for $y_1\neq y_2$, $d(\tau y_1,\tau y_2)$ is bounded away from 0. We thus have another application of Theorem 3.1:

THEOREM 3.3. *A subflow of a distal flow is distal, and the Ellis group of a subflow is the homomorphic image of the Ellis group of the flow.*

Given a flow (X,T) there is also defined a flow $(X\times X,T)$ by $\tau(x_1,x_2)=(\tau x_1,\tau x_2)$. The latter flow will not be minimal even if (X,T) is, since the diagonal is always a closed invariant subset. If G is the Ellis group of (X,T) then G is also the Ellis group of $(X\times X,T)$ where we define

$g(x_1, x_2) = (gx_1, gx_2)$. Similarly G is the Ellis group of (X^A, T) where X^A is any power of X.

4. Reduction of the problem. For the remainder of our discussion, (X, T) will denote a fixed distal flow. If (Y, T) is a quasi-isometric subflow of (X, T) then according to Definitions 2.4 and 2.5, there is a well ordered system (X_ξ, T) of subflows, for ordinals $\xi \leqq \eta$, with (X_0, T) the trivial flow, $(X_\eta, T) = (Y, T)$, and satisfying (b), (c) and (d) of Def. 2.4. Let us call such a system a *quasi-isometric system*. Every quasi-isometric system is supposed to have a last element.

Let H denote the set of all quasi-isometric systems in (X, T). Thus each $\theta \in H$ assigns to all the ordinals $\xi \leqq \eta(\theta)$ a subflow (X_ξ, T) such that the collection $\{(X_\xi, T), \xi \leqq \eta(\theta)\}$ forms a q. i. system. We impose the condition on the q. i. systems in H, that if $\xi < \xi'$, then $(X_{\xi'}, T)$ is a non-trivial extension of (X_ξ, T), that is, $(X_{\xi'}, T)$ is not isomorphic to (X_ξ, T). With this hypothesis, the ordinals occurring in a q. i. system correspond to cardinals not exceeding the cardinality of X. We thereby avoid difficulties of the meaningfulness of the set of all such system. We may define a partial ordering in H as follows. We say $\theta_1 \leqq \theta_2$ if $\eta(\theta_1) \leqq \eta(\theta_2)$ and if the flows (X_ξ, T) coincide for θ_1 and θ_2 whenever $\xi \leqq \eta(\theta_1)$. The system θ_1 is then to be a segment of θ_2.

LEMMA 4.1. *A totally ordered subset of H has an upper bound in H.*

By a totally ordered subset we mean one such that if θ_1 and θ_2 belong to the subset, either $\theta_1 \leqq \theta_2$ or $\theta_2 \leqq \theta_1$.

Proof. Let Σ be a totally ordered subset. If Σ has a largest member, this will be an upper bound for Σ. We may therefore suppose that the set of ordinals $\{\eta(\theta), \theta \in \Sigma\}$ has no largest member. Let η be the first ordinal larger than $\{\eta(\theta), \theta \in \Sigma\}$. If $\xi < \eta$ then $\xi < \eta(\theta)$ for some θ in Σ, so (X_ξ, T) is defined. Since Σ is totally ordered, the flow (X_ξ, T) is unambiguously defined. Let $(X_\xi, T) = \pi_\xi(X, T)$. Form the cartesian product $\prod_{\xi<\eta} X_\xi$ and let X_η be the subset of this product of elements of the form $(\pi_\xi(x))$, $x \in X$. The map $\pi_\eta : X \to X_\eta$ defined by $\pi_\eta(x) = (\pi_\xi(x))$ determines a subflow (X_η, T). For $\xi < \eta$, (X_ξ, T) is a subflow of (X_η, T) and (X_η, T) is a limit of $\{(X_\xi, T), \xi < \eta\}$. Now define a quasi-isometric system θ^* by taking the (X_ξ, T) for $\xi < \eta$, together with (X_η, T) as here defined. To verify that this is a q.i. system, it is only necessary to show that η is a limit ordinal, since the conditions (b), (c), (d) are satisfied for $\xi < \eta$. Suppose, then, that $\eta = \eta_0 + 1$. We would then have $\eta_0 = \eta(\theta_0)$ for some $\theta_0 \in \Sigma$, for otherwise η_0 could have been chosen in the place of η. But then θ_0 is clearly an upper bound, contrary to the assumption that Σ has no largest member. This proves the lemma.

PROPOSITION 4.1. *The set H possesses a maximal element.*

Proof. Zorn's lemma.

Using Proposition 4.1 we are able to bring about a considerable simplification in the problem of showing that every minimal distal flow is q.i. Namely we have

THEOREM 4.2. *Suppose (X,T) is a flow with the property that whenever (Z,T) is a proper subflow, then we may interpolate between (Z,T) and (X,T) a flow (Y,T), which is a subflow of (X,T) and a non-trivial isometric extension of (Z,T). Then (X,T) is quasi-isometric.*

Proof. If (X,T) is not quasi-isometric and θ is a maximal element of H, then the last member (Z,T) of the system θ is a proper subflow of (X,T). By hypothesis, there then exists an isometric extension (Y,T) of (Z,T) which is a subflow of (X,T), and $(Y,T)\neq(Z,T)$. Now let θ^* denote the q.i. system with (X_ξ,T) the same for θ^* as for θ, if $\xi\leqq\eta(\theta)$, and with $(X_{\eta(\theta)+1},T)=(Y,T)$. Then $\theta^*\geqq\theta$ and $\theta^*\neq\theta$, which contradicts the maximality of θ. This contradiction shows that the last member of θ must be (X,T) and so (X,T) is quasi-isometric.

The next six sections will be devoted to proving that if (X,T) is minimal distal, it fulfills the hypothesis of Theorem 4.2. We will thereby have proved Theorem 2.4 that every minimal distal flow is q.i. We conclude this section by giving an indication of the manner in which the proof will proceed. We suppose then that (Z,T) is a proper subflow of (X,T). We may think of Z as being an identification space of X; namely where the fibres of X over Z are identified as single points. If (Y,T) is to be interpolated between (X,T) and (Z,T), then Y is an identification space of X such that if two points of X are in the same fibre over Y they are necessarily in the same fibre over Z. Thus each fibre over Z may be decomposed into fibres over Y. This set of fibres over Y in a single fibre over Z constitutes the fibre of Y over Z. The space Y will be constructed by defining an equivalence relation in each fibre over Z and letting Y be the resulting identification space. If (Y,T) were given as an isometric extension of (Z,T), then there would exist a function $\rho(y_1,y_2)$ defined for y_1, y_2 in the same fibre over Z and satisfying the conditions of Definitions 2.1 and 2.2. The function ρ would induce a function R defined on pairs of points x_1, x_2 in the same fibre of X over Z and satisfying related conditions. What we shall do is to construct the function $R(x_1,x_2)$ first and retrieve the space Y from it by identifying x_1 and x_2 if $R(x_1,x_2)=0$. We notice that since $\rho(\tau y_1,\tau y_2)=\rho(y_1,y_2)$, we will also have $R(\tau x_1,\tau x_2)$

$=R(x_1, x_2)$, where $\tau \in T$. Since $R(x_1, x_2)$ is continuous, this will imply $R(gx_1, gx_2) = R(x_1, x_2)$ for g in the Ellis group of (X, T). The crux of the problem lies in showing that a non-constant continuous function $R(x_1, x_2)$ exists which is invariant under G. Now fix a fibre X_0 of X over Z, and let G_0 be the subgroup of G which takes X_0 into X_0. The restriction of R to pairs of points in X_0 defines a non-constant continuous function on $X_0 \times X_0$ invariant under G. It turns out that if we can find a function with these properties then we can also construct $R(x_1, x_2)$ for all the fibres. As a result, for a good part of the discussion, we shall confine our attention to the single fibre X_0.

5. The function $F(x, x')$ and the group G_0.

Definition 5.1. For $x, x' \in X$, define

$$(5.1) \qquad F(x, x') = \text{glb}\{d(\tau x, \tau x'), \tau \in T\} = \inf\{d(gx, gx'), g \in G\}$$

where G is the Ellis group of the distal flow (X, T).

The second equality in (5.1) follows from the fact that $d(gx, gx')$ is a continuous function on G, and T is dense in G. The following assertions follow directly from this definition.

PROPOSITION 5.1. (a) $F(x, x') = F(x', x)$, (b) $F(x, x') \geqq 0$ and equality holds only if $x = x'$, (c) $F(gx, gx') = F(x, x')$ for $g \in G$, (d) $F(x, x')$ is an upper semi-continuous function on $X \times X$.

Proof. (a), (b), and (c) are obvious under the hypothesis that (X, T) is distal. (d) follows from the fact that $F(x, x')$ is a greatest lower bound of a family of continuous functions.

We remark that if the function $F(x, x')$ turns out to be continuous, then it is possible to show that (X, T) is actually equicontinuous. Thus in the general case, we cannot improve (d).

LEMMA 5.1. *Let $F(x, x') < a$. There exists $\epsilon > 0$ such that whenever $F(x', x'') < \epsilon$, then $F(x, x'') < a$.*

Proof. Let T_n denote an increasing sequence of compact subsets of T such that $\cup T_n = T$. Consider now the orbit $G(x, x') = \{(gx, gx'), g \in G\} \subset X \times X$. Suppose $F(x, x') < b < a$. The subset of pairs $(u, u') \in G(x, x')$ such that $d(u, u') < b$, forms a relatively open set $U \subset G(x, x')$, and since $F(x, x') < b$, $U \neq \emptyset$. We claim that there exists an n such that for any $(u, u') \in G(x, x')$, $T_n(u, u') \cap U \neq \emptyset$. Suppose this were not the case. Then for every n we could find $(u_n, u'_n) \in G(x, x')$ with $T_n(u_n, u'_n) \subset G(x, x') - U$ which is closed. So for $m > n$, $T_n(u_m, u'_m) \subset G(x, x') - U$ since $T_n \subset T_m$.

For some subsequence we will have $(u_{n_k}, u'_{n_k}) \to (v, v')$, and so $T_n(v,v') \subset G(x,x') - U$. This is true for all n, so $T(v,v') \cap U = \emptyset$, so $G(v,v') \cap U = \emptyset$ which is impossible since $U \subset G(v,v')$ and $U \neq \emptyset$. This proves the existence of n as asserted.

Now recall that the mapping $(\tau, x) \to \tau x$ is continuous from $T \times X \to X$. This implies that for τ in the compact set $T_n \subset T$, the maps $\tau: X \to X$ are equicontinuous. So there exists an $\epsilon > 0$ such that if $d(x_1, x_2) < \epsilon$, then $d(\tau x_1, \tau x_2) < a - b$ for all $\tau \in T_n$. Suppose now that $F(x', x'') < \epsilon$. There exists $\tau_0 \in T$ such that $d(\tau_0 x', \tau_0 x'') < \epsilon$, and hence $d(\tau\tau_0 x', \tau\tau_0 x'') < a - b$ for $\tau \in T_n$. By the preceding paragraph, since $(\tau_0 x, \tau_0 x') \in G(x,x')$, there exists $\tau_1 \in T_n$ with $(\tau_1\tau_0 x, \tau_1\tau_0 x') \in U$ whence $d(\tau_1\tau_0 x, \tau_1\tau_0 x') < b$, and so $d(\tau_1\tau_0 x, \tau_1\tau_0 x'') < a$. This proves that $F(x, x'') < a$ as required.

PROPOSITION 5.2. *The sets $U_a(x) = \{x' : F(x,x') < a\}$ form a basis of topology in X.*

Proof. It must be shown that if $U_a(x_1) \cap U_b(x_2) \neq \emptyset$, then this intersection is the union of sets of this kind. But if $x' \in U_a(x_1)$ then $U_{\epsilon_1}(x') \subset U_a(x_1)$ by Lemma 5.1. Similarly if $x' \in U_b(x_2)$, $U_{\epsilon_2}(x') \subset U_b(x_2)$. So if $\epsilon \leqq \epsilon_1$ and $\epsilon \leqq \epsilon_2$ $U_\epsilon(x') \subset U_a(x_1) \cap U_b(x_2)$.

From this point on through § 10, (Z, T) will denote a given subflow of (X, T), z_0 will be a point in Z, and X_0 will denote the fibre $X_{z_0} = \pi^{-1}(z_0)$ where $(Z, T) = \pi(X, T)$. These will be fixed throughout the discussion. Now let $G_0 = \{g \in G : gX_0 \subset X_0\}$.

LEMMA 5.2. *Let $x, x' \in X_0$ and $g \in G$. If $gx = x'$ then $g \in G_0$.*

Proof. If $x_1 \in X_0$, then $\pi(x_1) = \pi(x) = z_0$. Hence $\pi(gx_1) = \pi(gx) = \pi(x') = z_0$ by Lemma 3.1. Therefore $gx_1 \in X_0$ and so $gX_0 \subset X_0$.

PROPOSITION 5.3. *G_0 is a closed subgroup of G and G_0 is transitive on X_0.*

Proof. G_0 is clearly closed and a semigroup. It is a group since if $gx = x'$, where $x, x' \in X_0$, then $g^{-1}x' = x$, and by the preceding lemma, $g^{-1} \in G_0$. Finally G_0 is transitive on X_0 since given any $x, x' \in X_0$ there is some $g \in G$ with $gx = x'$ (G is transitive on X since (X, T) is minimal) and by the preceding lemma, $g \in G_0$.

Definition 5.2. The relative topology of X_0 as a subset of X, and the relative topology of G_0 as a subset of G will be referred to as the *E-topologies* on these spaces.

Under the E-topology X_0 and G_0 are compact Hausdorff spaces.

Definition 5.3. The topology of X_0 determined by the basis $\{U_a(x) \cap X_0\}$ (see Prop. 5.2) will be referred to as the *F-topology* on X_0.

Definition 5.4. By the *F-topology* on G_0 we shall mean the topology determined by the sub-basis of sets of the form $W_a(x,x') = \{g \in G_0 : F(x,gx') < a\}$ where $x, x' \in X_0$.

PROPOSITION 5.4. a) *The F-topology on X_0 is weaker (i. e. no stronger) than the E-topology.*

b) *The transformations of G_0 on X_0 are continuous with respect to the F-topology.*

c) *The function $F(x,x')$ as a function of x' is upper semi-continuous with respect to the F-topology on X_0.*

Proof. a) Since $F(x,x')$ is upper semi-continuous relative to the E-topology on X_0, $U_a(x) = \{x' : F(x,x') < a\}$ is E-open.

b) follows from the fact that $F(gx,gx') = F(x,x')$ whence $gU_a(x) = U_a(gx)$.

c) is a restatement of Lemma 5.1.

As a conesquence of (b) we see that if the E- and F-topologies coincided on X_0, then G_0 would act on X_0 by continuous transformations. The later developments will show that this would imply that (X,T) is already an isometric extension of (Z,T). Thus in general the F-topology on X_0 is strictly weaker than the E-topology. Since the latter is compact, we conclude that the former cannot be Hausdorff.

PROPOSITION 5.5. (a) *the F-topology on G_0 is weaker than the E-topology.*

b) *the maps $g \to g\gamma$, $g \to \gamma g$, for fixed $\gamma \in G_0$, and the map $g \to g^{-1}$ are all continuous relative to the F-topology.*

c) *For fixed $x \in X_0$, the map $g \to gx$ of $G_0 \to X_0$ is continuous relative to the F-topologies on the two spaces.*

Proof. (a) follows from the fact that $F(x,gx') \leq d(x,gx')$.

(b) is a consequence of

$$\begin{aligned} W_a(x,x')\gamma &= W_a(x,\gamma^{-1}x') \\ \gamma W_a(x,x') &= W_a(\gamma x,x') \qquad (5.1) \\ W_a(x,x')^{-1} &= W_a(x',x), \end{aligned}$$

all of which are consequences of the invariance of $F(x,x')$ under G.

(c) $\{g : gx \in U_a(x')\} = \{g : F(x',gx) < a\} = W_a(x',x)$.

If we take (Z,T) in this discussion to be the trivial flow, then $X_0 = X$ and $G_0 = G$. The topologies introduced are then topologies for the original space and the entire Ellis group. It would have been more straightforward to introduce the F-topologies on these. However in our discussion it will be necessary to treat the fiber X_0 (when (Z,T) is not trivial) as the total space, and for this reason we have introduced these topologies directly for X_0 and G_0.

6. The discontinuities of $F(x,\cdot)$. Set $F_x(x') = F(x,x')$. Here $x, x' \in X_0$ and F_x is a function on X_0. With respect to the E-topology (metric topology) on X_0, F_x is an upper semi-continuous function (Prop. 5.1 (d)) on X_0 and hence its discontinuities form a set of the first category. Here the discontinuities are with respect to the E-topology. These form a subset of the discontinuities with respect to the F-topology (which is weaker). Now by Proposition 5.4 (c), F_x is still an upper semi-continuous function relative to the F-topology so that the set of discontinuities with respect to the F-topology form a set of first category (i.e. countable union of closed nowhere dense sets) with respect to the F-topology. However it is not clear that for the F-topology (which is not metric) the first category sets are "small"—for example, it is not clear that the whole space X_0 is not of the first category. We shall show in this section that the two notions of first category in fact coincide.

LEMMA 6.1. *Let Δ be an F-closed subset of X_0 which contains an E-open set. Then Δ contains an F-open set.*

Proof. Let x_0 be any point in X_0 and set $\Gamma = \{g \in G_0 : gx_0 \in \Delta\}$. If Δ_0 is an E-open set contained in Δ and $\Gamma_0 = \{g \in G_0 : gx_0 \in \Delta_0\}$, then Γ_0 is an E-open subset of Γ in G_0 (§3). Now the map $g \to g\gamma$ is continuous on G_0, for $\gamma \in G_0$, in the E-topology of G_0. Hence $\Gamma_0\gamma$ is E-open for any $\gamma \in G_0$. So $\bigcup_{\gamma \in G_0} \Gamma_0\gamma$ is an open covering of G_0. Since G_0 is compact (in both topologies) there is a finite set $\{\gamma_1, \cdots, \gamma_r\}$ with $\bigcup \Gamma_0\gamma_i = G_0$ and *a fortiori* $\bigcup \Gamma\gamma_i = G_0$. By Prop. 5.5 (c), the map $g \to gx_0$ of $G_0 \to X_0$ is continuous with respect to the F-topologies, and since Δ is F-closed, Γ will be F-closed. By Prop. 5.5 (b), the sets $\Gamma\gamma_i$ are also F-closed. Now let s be the smallest integer for which $G_0 = \bigcup_{i=1}^{s} \Gamma\gamma_i$. Then $G_0 - \bigcup_{i=1}^{s-1} \Gamma\gamma_i$ is non-empty and F-open in G_0. Hence $\Gamma\gamma_s$ contains an F-open set, and so Γ contains an F-open set, Γ_{00}. Since multiplication on the left is continuous for the F-topology on G_0, it follows that $\bigcup_{\gamma \in G_0} \gamma\Gamma_{00}$ is an open covering of G_0. Again by appropriate choice of

$\{\gamma'_1, \cdots, \gamma'_m\}$, $\bigcup \gamma'_i \Gamma_{00} = G_0$, and $\bigcup \gamma'_i \Gamma = G_0$. Since G_0 is transitive on X_0 (Prop. 5.3) we have $\Gamma x_0 = \Delta$ and $\bigcup \gamma'_i \Delta = X_0$. Now let l be the least integer for which $\bigcup_{i=1} \gamma'_i \Delta = X_0$. As before, since Δ is F-closed and therefore each $\gamma'_i \Delta$ is F-closed (Prop. 5.4 (b)), we conclude that $\gamma'_l \Delta$, and therefore Δ, contains an F-open set. This concludes the proof of the lemma.

PROPOSITION 6.1. *If a subset of X_0 is of the first category with respect to the F-topology then it is of first category with respect to the E-topology.*

Proof. Suppose Δ is a nowhere dense set relative to the F-topology. Let Δ_E and Δ_F denote respectively the E- and F-closures of Δ. Then $\Delta_E \subset \Delta_F$. If Δ_E contained an E-open set then Δ_F would contain an E-open set and by Lemma 6.1, Δ_F would contain an F-open set contrary to the hypothesis that Δ was nowhere dense in the F-topology. Hence Δ is nowhere dense in the E-topology and this proves the proposition.

In the sequel we shall use the phrase "of first category" to refer to the (metric) E-topology of X_0. (We have not shown that any such set is also of first category in the F-topology but this will not be needed).

We are now interested in showing that if $\phi(x)$ is an upper semi-continuous function on X_0 with respect to the F-topology, then its F-discontinuities form a set of first category. By Prop. 6.1, this will follow once it is shown that the set of F-discontinuities is of first category relative to the F-topology. This is true quite generally and is proven just as in the metric case. For lack of a reference we present the proof here.

Definition 6.1. Let ϕ be an upper semi-continuous function on X_0 relative to the F-topology. We say that ϕ has an F-discontinuity c at x_0, if for any $\epsilon > 0$, there exists x with $F(x_0, x) < \epsilon$ and $\phi(x) < \phi(x_0) - c$.

If for no $c > 0$ does ϕ have an F-discontinuity c at x_0, then since ϕ is upper semi-continuous, $\lim_{F(x, x_0) \to 0} \phi(x) = \phi(x_0)$ and ϕ is F-continuous at x.

LEMMA 6.2. *If ϕ is bounded and upper semi-continuous in the F-topology, the set of points at which ϕ has F-discontinuity $c > 0$ is nowhere dense in the F-topology.*

Proof. Denote this set by $\Delta(c)$. It is easy to see that the F-closure of $\Delta(c)$ is contained in $\Delta(c - \epsilon)$ for any $\epsilon > 0$. It suffices therefore to show that $\Delta(c)$ contains no F-open set. Suppose U is an F-open set contained in $\Delta(c)$. Let $\alpha = \inf\{\phi(x), x \in U\}$. Then there exists $x_1 \in U$ with $\phi(x_1) < \alpha + c/2$. Moreover in every F-neighborhood of x_1 there are points x_2 with $\phi(x_2) < \phi(x_1) - c < \alpha - c/2$. But some F-neighborhood of x_1 is contained in U and this contradicts the choice of α. This proves the lemma.

The set of all F-discontinuities of ϕ is $\bigcup_{m=1}^{\infty} \Delta(1/m)$, and so is of the first category relative to F, and by Prop. 6.1, it is of first category. In particular

PROPOSITION 6.2. *For each x, the function F_x on X_0 is continuous with respect to the F-topology at all points but for a set of the first category.*

PROPOSITION 6.3. *Let $x' \in X_0$. The set of x such that F_x has an F-discontinuity at x' is of the first category in X_0.*

Proof. This set is the union of those x for which F_x has an F-discontinuity $1/m$ at x', for $m = 1, 2, 3, \cdots$. It will suffice to show that each of these sets is nowhere dense (in the E-topology). Let Δ' be the set of x such that F_x has an F-discontinuity c at x', and let Δ be the set of x such that for every $\epsilon > 0$, F_x has an F-discontinuity $c - \epsilon$ at x'. $\Delta' \subset \Delta$, and so it suffices to show that Δ is nowhere dense. Let $\Gamma = \{g \in G_0 : gx' \in \Delta\}$. Since each $g \in G_0$ is a homeomorphism of X_0, and $F(x, x')$ is invariant under G_0, it follows that if $F_{gx'}$ has F-discontinuity $c - \epsilon$ at x', then $F_{x'}$ has an F-discontinuity $c - \epsilon$ at $g^{-1}x'$, and conversely. So $\Gamma^{-1} = \{g \in C_0 : gx' \in \Delta^*\}$ where Δ^* is the set of points at which $F_{x'}$ has an F-discontinuity $c - \epsilon$ for every $\epsilon > 0$. It is easily seen that Δ^* is F-closed and by Lemma 6.2, Δ^* contains no F-open set. Since Δ^* is F-closed, it follows by Prop. 5.5 that Γ^{-1} is F-closed and therefore also Γ is F-closed. Therefore Γ is E-closed, and so $\Delta = \Gamma x'$ is E-closed (but not necessarily F-closed, since an F-compact set need not be F-closed, the F-topology on X_0 not being Hausdorff). Suppose Δ contained an E-open set, then Γ would also contain an E-open set Γ_0. Now the map $g \to g\gamma$ is continuous on G_0 in the E-topology, so the sets $\{\Gamma_0\gamma\}$ are open and cover G_0 whence $G_0 = \bigcup \Gamma\gamma_i$ for a finite set $\{\gamma_i\}$ then $G_0 = \bigcup \gamma_i^{-1}\Gamma^{-1}$, and $X_0 = \bigcup \gamma_i^{-1}\Delta^*$. But this is impossible since the $\gamma_i^{-1}\Delta^*$ are F-closed and contain no F-open sets. This shows that Δ and a fortiori Δ' is nowhere dense and this proves the proposition.

Definition 6.2. If a statement is valid for all $x \in X_0$ outside of a set of the first category, we will say it holds *almost everywhere* (a.e.).

The main purpose of the foregoing discussion is to obtain the following result.

THEOREM 6.1. *If $x_n \to x_0$ in the F-topology, then $F_{x_n}(x') \to F_{x_0}(x')$ almost everywhere.*

Proof. The set of x' such that x_0 is not a point of F-continuity of $F_{x'}$ is of first category. Outside this set $F_{x'}(x_n) \to F_{x'}(x)$. But this is the same as $F_{x_n}(x') \to F_{x_0}(x')$. (Prop. 5.1 (a)).

7. The homogeneous space M.

Definition 7.1. For $x_1, x_2 \in X_0$, write $x_1 \sim x_2$ in case $F_{x_1}(x') = F_{x_2}(x')$ almost everywhere in X_0. Let M denote the space of equivalence classes of X_0 with respect to this equivalence relation.

That $\sim$ is an equivalence relation is clear. We would like to show that the transformations of G_0 preserve this equivalence relation.

LEMMA 7.1. *If $x_1 \sim x_2$ then $F_{x_1}(x') = F_{x_2}(x')$ at all x' which are points of E-continuity of both F_{x_1} and F_{x_2} (and a fortiori points of F-continuity). Conversely if $F_{x_1}(x') = F_{x_2}(x')$ at points of F-continuity then $x_1 \sim x_2$.*

Proof. The set of x' for which $F_{x_1}(x') = F_{x_2}(x')$ is valid is E-dense if $x_1 \sim x_2$. So this equality must hold at points of E-continuity by passage to the limit. For the converse we refer to Prop. 6.2 according to which F_{x_1} and F_{x_2} are simultaneously continuous a. e.

PROPOSITION 7.1. *If $x_1 \sim x_2$, then $gx_1 \sim gx_2$ for $g \in G_0$.*

Proof. If $F(x_1, x') = F(x_2, x')$ whenever x' is a point of F-continuity of F_{x_1} and F_{x_2}, then $F(gx_1, gx') = F(gx, gx')$, and g being a homeomorphism of X_0 with repsect to the F-topology, the points of F-continuity of F_{gx_i} are the images under g of the points of F-continuity of F_{x_i}. The proposition then follows from Lemma 7.1.

Let π_0 denote the natural map from X_0 onto M taking each point into its equivalence class. M becomes a topological space by designating as open the sets V for which $\pi_0^{-1}(V)$ is an F-open set in X_0. Then π_0 is clearly continuous. Although X_0 was not Hausdorff in the F-topology we have

THEOREM 7.2. *M is a compact Hausdorff space.*

Proof. As the continuous image of X_0, M is compact. Now let y_1 and y_2 be distinct points of M and let $\Delta_1 = \pi_0^{-1}(y_1)$, $\Delta_2 = \pi_0^{-1}(y_2)$. Let $x_1 \in \Delta_1$, $x_2 \in \Delta_2$, and form the F-neighborhoods $U_\epsilon(x_1)$ and $U_\epsilon(x_2)$ ($U_\epsilon(x_i) = \{x : F(x_i, x) < \epsilon\}$). Suppose these sets overlapped for every $\epsilon > 0$. We could then find a sequence $u_n \in U_{1/n}(x_1) \cap U_{1/n}(x_2)$ so that $u_n \to x_1$ and $u_n \to x_2$ in the sense of the F-topology on X_0. But then, by Theorem 6.4, $F_{u_n} \to F_{x_1}$ and $F_{u_n} \to F_{x_2}$ a. e. so that $F_{x_1} = F_{x_2}$ a. e. and $x_1 \sim x_2$. This would imply $y_1 = \pi_0(x_1) = \pi_0(x_2) = y_2$ contrary to our hypothesis. The conclusion is that any point of Δ_1 can be separated from any point of Δ_2 by non-overlapping F-open sets.

Next we notice that Δ_1 and Δ_2 are compact. For suppose $x'_n \sim x_1$ and $x'_n \to x'$ in the F-topology. Then $F_{x'_n} \to F_{x'}$ a. e. But $F_{x'_n} = F_{x_1}$ a. e. whence

$F_{x'} = F_{x_1}$ a. e. and $x' \sim x_1$, $x' \in \Delta_1$. So Δ_1 is closed and hence compact. The same holds for Δ_2. Since we can separate any pair of points from Δ_1 and Δ_2 by non-overlapping F-open sets, a standard argument shows that there exist F-open sets U_1 and U_2 with $\Delta_1 \subset U_1$, $\Delta_2 \subset U_2$ and $U_1 \cap U_2 = \emptyset$. (The same argument used to prove that a compact Hausdorff space is normal). Let U'_1 denote the set of points not equivalent ($\sim$) to any point outside U_1 and define U'_2 similarly. Then $\Delta_1 \subset U'_1 \subset U_1$ and $\Delta_2 \subset U'_2 \subset U_2$. We claim that U'_1 and U'_2 are F-open. Suppose $u_n \notin U'_1$ and $u_n \to u$ in the F-topology. There will exist $v_n \notin U_1$ with $u_n \sim v_n$. Passing to a subsequence of the original sequence we may suppose that v_n converges in the F-topology, say to v. (In fact if $v_n \to v$ in the E-topology, then $v_n \to v$ in the F-topology). We then have $F_{u_n} \to F_u$ a. e. and $F_{v_n} \to F_v$ a. e. and $F_{u_n} = F_{v_n}$ a. e., whence $F_u = F_v$ a. e. and $u \sim v$. But $v \notin U_1$ so $u \notin U'_1$, and this proves that U'_1 is F-open. The same argument shows that U'_2 is F-open. Now any point equivalent to a point of U'_1 is clearly in U'_1. We find then that $\pi_0^{-1}\pi_0(U'_1) = U'_1$ and similarly $\pi_0^{-1}\pi_0(U'_2) = U'_2$. This implies that $\pi_0(U'_1)$ and $\pi_0(U'_2)$ are disjoint open sets in M. Finally $y_1 = \pi_0(x_1) \in \pi_0(U'_1)$ and $y_2 \in \pi_0(U'_2)$ which proves that M is Hausdorff.

Definition 7.2. For $g \in G_0$, let $\theta(g)$ denote the transformation of M defined by

$$\theta(g)\pi_0(x) = \pi_0(gx). \tag{7.1}$$

where $x \in X_0$.

This definition is unambiguous since by Prop. 7.1, if $\pi_0 x_1 = \pi_0 x_2$, also $\pi_0 g x_1 = \pi_0 g x_2$.

Definition 7.3. Let H be the group $\theta(G_0)$. We define a topology on H by designating as open the sets W for which $\theta^{-1}(W)$ is F-open in G_0.

PROPOSITION 7.3. *If $h_0 \in H$, the operations $h \to hh_0$ and $h \to h_0 h$ are continuous on H. If $y \in M$, the map $h \to hy$ is continuous from H to M. Finally, the group H is compact and Hausdorff.*

Proof. Let W be an open set in H and set $h_0 = \theta(g_0)$. Now $\theta(g)h_0 \in W$ is equivalent to $\theta(g)\theta(g_0) \in W$ or $\theta(gg_0) \in W$, $gg_0 \in \theta^{-1}(W)$. The latter set is F-open and so is $\theta^{-1}(W)g_0^{-1}$. This shows that $\theta^{-1}\{h : hh_0 \in W\} = \theta^{-1}(W)g_0^{-1}$ is F-open and so $\{h : hh_0 \in W\}$ is open and $h \to hh_0$ is continuous. A similar argument applies to $h \to h_0 h$. We have used here the fact that both left and right multiplications are continuous for the F-topology on G_0. Now let V be open in M and let $y = \pi_0(x)$. $\{g : \theta(g)\pi_0(x) \in V\} = \{g : \pi_0(gx) \in V\}$

$=\{g: gx \in \pi_0^{-1}(V)\}$. Since $\pi_0^{-1}(V)$ is F-open in X_0, this set of g is F-open in G_0. Hence $\theta^{-1}\{h: hy \in V\}$ is F-open in G_0 and so $\{h: hy \in V\}$ is open in H and $h \to hy$ is continuous. Since H is the continuous image of G_0 it is compact. To prove it is Hausdorff suppose for a pair of elements in H, $h_1 \neq h_2$. There exists a point y_0 such that $h_1y_0 \neq h_2y_0$. Since M is Hausdorff we can find two open sets V_1 and V_2 in M, with $h_1y_0 \in V_1$, $h_2y_0 \in V_2$ and $V_1 \cap V_2 = \emptyset$. Set $W_1 = \{h: hy_0 \in V_1\}$, $W_2 = \{h: hy_0 \in V_2\}$. Since $h \to hy_0$ is continuous W_1 and W_2 are open in H. Clearly $h_1 \in W_1$, $h_2 \in W_2$ and $W_1 \cap W_2 = \emptyset$. This completes the proof of the proposition.

The following result is due to Ellis ([5]).

PROPOSITION 7.4. *Let H be a group which is provided with a topology for which H is compact, Hausdorff and for which left and right multiplication are continuous. Then H is a topological group; i.e., the map $(h_1, h_2) \to h_1h_2$ is continuous from $H \times H \to H$. More generally, if H operates on a compact Hausdorff space M in such a way that for $h \in H$, $y \in M$, hy is separately continuous in h and in y, then hy is jointly continuous; i.e., $(h, y) \to hy$ is a continuous map from $H \times M \to M$.*

Since the hypotheses of this proposition are verified in our case of Prop. 7.3, we conclude that H is a compact topological group the map $(h, y) \to hy$ is continuous from $H \times M$ to M. Now, since X_0 has a countable base for its topology so does M, and being compact Hausdorff, M is metrizable. Let $D'(y_1, y_2)$ denote a metric. We claim that

$$D(y_1, y_2) = \sup_{h \in H} D'(hy_1, hy_2)$$

is an equivalent metric on M. Clearly it is a metric; to show it is equivalent we need to show that for $y_0 \in M$ and $\epsilon > 0$, there is a neighborhood V of y_0 with $D(y, y_0) < \epsilon$ for $y \in V$. Now by what we have shown, $D'(hy, hy_0)$ is a continuous function of the pair $(h, y) \in H \times M$. It vanishes on $H \times y_0$, so there is a finite covering of $H \times y_0$ by open sets of the form $W_i \times V_i$ on which this function $< \epsilon/2$. Take $V = \cap V_i$, then for $y \in V$, $D'(hy, hy_0) < \epsilon/2$ for all h and so $D(y, y_0) \leqq \epsilon/2 < \epsilon$. This shows that D may be taken as the metric on M. Since $D(y_1, y_2) = D(hy_1, hy_2)$ we have

THEOREM 7.5. *M is a homogeneous compact metric space and H acts transitively on M by isometries.*

Proof. We need only remark that H is transitive on M because G_0 is transitive on X_0. Hence M is homogeneous.

THEOREM 7.6. *M will not reduce to a single point unless the fibre X_0 consisted of just one point.*

Proof. If M consisted of one point, then $x_1 \sim x_2$ for all pairs of points in X_0, and $F(x_1, x) = F(x_2, x)$ a.e. Now let $x \to x_1$ while ranging over the simultaneous points of continuity of F_{x_1} and F_{x_2}. Then $\lim F(x_2, x) = 0$. If, in addition, x_1 is a point of continuity of F_{x_2}, then $\lim F(x_2, x) = F(x_2, x_1)$, so that $F(x_2, x_1) = 0$, or $x_1 = x_2$. Since almost all points of X_0 are points of continuity of F_{x_2}, this implies that $X_0 - \{x_2\}$ is a set of the first category. This can only take place if X_0 reduces to a single point.

Definition 7.4. By a *semi-metric* on a space Ω we mean a function $\sigma(w_1, w_2)$ on $\Omega \times \Omega$ satisfying

(i) $\sigma(w_1, w_2) \geqq 0$, (ii) $\sigma(w_1, w_3) \leqq \sigma(w_1, w_2) + \sigma(w_2, w_3)$,

(iii) $\sigma(w_1, w_2) = \sigma(w_2, w_1)$, (iv) $\sigma(w_1, w_1) = 0$.

(σ is a metric except that it may vanish for distinct points).

Definition 7.5. Let $\sigma(x_1, x_2)$ be given by

$$\sigma(x_1, x_2) = D(\pi_0(x_1), \pi_0(x_2)).$$

Insofar as they will be applied in the sequel, the results of the last two sections may be summarized by the following theorem.

THEOREM 7.7. *If (X, T) is a non-trivial extension of (Z, T) and X_0 is a fibre of X with respect to Z, G_0 is the subgroup of the Ellis group taking X_0 into itself, then there exists a non-constant function $\sigma(x_1, x_2)$ on $X_0 \times X_0$ such that*

(a) *$\sigma(x_1, x_2)$ is a semi-metric on X_0,*

(b) *$\sigma(gx_1, gx_2) = \sigma(x_1, x_2)$ for $g \in G_0$,*

(c) *$\sigma(x_1, x_2)$ is continuous with respect to the F-topology on X_0; more precisely*

(d) *for any $\epsilon > 0$, there is a δ such that whenever $F(x_1, x_2) < \delta$, $\sigma(x_1, x_2) < \epsilon$.*

Proof. (a) and (b) follow directly from the definition of σ. (c) follows from (d) so we proceed to prove the latter. Suppose (d) were not valid and there existed pairs $(x_1^{(n)}, x_2^{(n)})$ with $F(x_1^{(n)}, x_2^{(n)}) \to 0$ and $\sigma(x_1^{(n)}, x_2^{(n)}) \geqq \epsilon$. Both F and σ are invariant under G_0; applying an appropriate $g^{(n)}$ to $x_1^{(n)}$ and $x_2^{(n)}$ we may assume that $x_2^{(n)}$ is independent of n, say $x_2^{(n)} = x_2$. But then $x_1^{(n)} \to x_2$ in the F-topology and so $\pi_0(x_1^{(n)}) \to \pi_0(x_2)$ in M and $\sigma(x_1^{(n)}, x_2)$ $= D(\pi_0(x_1^{(n)}), \pi_0(x_2))$ tends to 0. This completes the proof of the theorem.

It might be pointed out that if (Z,T) is the trivial flow, then our argument could stop at this point. X_0 would then be X and M would be an identication space of X, and we would have that (M,T) is a subflow of (X,T). Since H acts on M by isometries, T does so also. We have thus shown that every minimal distal flow has a non-trivial equicontinuous subflow. Our purpose in the next three sections is to extend what we have found for the single fibre X_0 to the entire space X.

8. The map π is open. In this section we prove a lemma which implies that if (X,T) is minimal distal and $(Z,T)=\pi(X,T)$ is any subflow, then the map $\pi\colon X\to Z$ is an open map. This will be used repeatedly in the subsequent sections.

LEMMA 8.1. *For any $\epsilon>0$ and $z\in Z$, there exists a finite subset $\Sigma\subset X_z(=\pi^{-1}(z))$ such that for each g in the Ellis group of (X,T), the finite set $g\Sigma$ is ϵ-dense in gX_z. (That is, every point in gX_z is at a distance $<\epsilon$ from some point of $g\Sigma$.)*

Proof. Such a set could certainly be chosen so that this condition be satisfied for one g; just take $\Sigma=g^{-1}\Sigma'$ where Σ' is an ϵ-dense subset of gX_z. (Recall Lemma 3.1 according to which g takes fibres over Z with fibres over Z.) For each g, then, let $\Sigma(g)$ denote such a set, but for $\epsilon/3$ instead of ϵ. Let x_0 be any point in X_z and set $W_\delta(g_1)=\{g\colon d(gx_0,g_1x_0)<\delta\}$ where g_1 is some fixed member of G. W_δ is then an E-open neighborhood of g_1. We claim that for sufficiently small δ, every point of gX_z is within a distance $\epsilon/3$ of a point of g_1X_z if $g\in W_\delta(g_1)$. If this were not the case we could find a sequence $x'_n\in g'_nX_z$ with $d(g'_nx_0,g_1x_0)\to 0$ and x'_n no closer than $\epsilon/3$ to any point of g_1X_z. Passing to a subsequence we may suppose that $x'_n\to x'$. Then $x'\notin g_1X_z$. Now let π be the map from X to Z. Since $x'_n\in g'_nX_z$, $\pi(x'_n)=\pi(g_1x_0)\to\pi(g_1x_0)$. Since $x'_n\to x'$, $\pi(x'_n)\to\pi(x')$ so $\pi(x')=\pi(g_1x_0)$ which implies that $x'\in g_1X_z$. This is a contradiction, and so for each g_1, δ exists as asserted. (δ may depend on g_1). Next let

$$W'(g_1)=\{g\colon d(gu,g_1u)<\epsilon/3 \text{ for every } u\in\Sigma(g_1)\}.$$

$W'(g_1)$ is again an E-open set. Take $g\in W'(g_1)\cap W_\delta(g_1)$ where δ is chosen as above. If $x\in gX_z$ then x is within $\epsilon/3$ of some point x_1 in g_1X_z. x_1 in turn is within $\epsilon/3$ of a point g_1u where $u\in\Sigma(g_1)$. Finally g_1u is within $\epsilon/3$ of gu, so x is within ϵ of gu. This shows that $\Sigma(g_1)$ fulfills the condition of the lemma for an E-open set about g_0. These open sets cover G, so a finite subset will do, say the sets corresponding to $\Sigma(g_i)$, $i=1,\cdots,m$. If we take $\Sigma=\cup\,\Sigma(g_i)$, then for every g, $g\Sigma$ will be ϵ-dense in gX_z.

THEOREM 8.1. *If (X,T) is distal and (Z,T) is minimal where $(Z,T)=\pi(X,T)$ then the map $\pi\colon X\to Z$ is open.*

Proof. Let V be an open set in X, we wish to prove that $\pi(V)$ is open. If this were not the case we would have a sequence $\{z_n\}$ in Z with $z_n\to z\in\pi(V)$ and $z_n\notin\pi(V)$. Let $z=\pi(x)$, $x\in V$ and suppose that V contains an ϵ-neighborhood of x. Since (Z,T) is minimal, G is transitive on the fibres of X (Corollary to Theorem 3.2). We may therefore write $z_n=\pi(g_nx_0)$ where x_0 is some point of X. Now let X_0 be the fibre passing through x_0, $X_0=\pi^{-1}\pi(x_0)$, and let Σ be a finite subset of X_0 with $g\Sigma$ ϵ-dence in gX_0 for every g. Passing to a subsequence, if necessary, we may suppose that g_nu converges for each $u\in\Sigma$. The limits may be written as gu for some $g\in G$ (Σ corresponds to a point in a product X^m, $\Sigma=(u_1,\cdots,u_m)$, and $G(u_1,\cdots,u_m)$ is compact). Now if $u\in\Sigma$, $\pi(g_nu)=\pi(g_nx_0)$ since $\pi(u)=\pi(x_0)$, so $\pi(g_nu)=z_n\to z$. Therefore $\pi(gu)=z$. Since $\pi(x)=z$ and $gu\in gX_0$, we have $x\in gX_0$. Hence x is within ϵ of some gu, $u\in\Sigma$. Since $g_nu\to gu$, it follows that for n sufficiently large, g_nu is within ϵ of x; hence $g_nu\in V$. But then $z_n=\pi(g_nu)\in\pi(V)$ contrary to our assumption. This shows that π is open.

COROLLARY. *If $\{z_n\}$ is a sequence in Z with $z_n\to z$ and $\pi(x)=z$, then there exists a sequence $\{x_n\}$ with $\pi(x_n)=z_n$ and $x_n\to x$.*

Proof. Let $\epsilon_k\to 0$, and let V_k be an ϵ_k-neighborhood of x. Since $\pi(V_k)$ is open, there is an N_k with $z_n\in\pi(V_k)$ for $n>N_k$. Here we may suppose that N_k is an increasing sequence. Then for $N_k<n\leqq N_{k+1}$ let $x_n\in V_k\cap\pi^{-1}(z_n)$. Clearly $x_n\to x$ and $\pi(x_n)=z_n$.

9. The function $R(x_1,x_2)$.

Definition 9.1. *If* $x_1,x_2\in X$ are in the same fibre over Z, i.e. $\pi(x_1)=\pi(x_2)$, then set $S(x_1,x_2)=\sigma(gx_1,gx_2)$ where σ is the function on $X_0\times X_0$ of Definition 7.5, and $g\in G$ is such that $gx_1\in X_0$.

Since (X,T) is minimal, such a g exists. Now if $gx_1\in X_0$ and $g'x_1\in X_0$, then $g'g^{-1}$ takes a point of X_0 into a point of X_0. By Lemma 5.2, $g'g^{-1}\in G_0$. Hence, using Theorem 7.7 (b),

$$\sigma(gx_1,gx_2)=\sigma(g'g^{-1}gx_1,g'g^{-1}gx_2)=\sigma(g'x_1,g'x_2).$$

This shows that Def. 9.1 is unambiguous.

Definition 9.2. For $x_1,x_2\in X$, $\pi(x_1)=\pi(x_2)$ let

$$R(x_1,x_2)=\limsup_{(u_1,u_2)\to(x_1,x_2)} S(u_1,u_2). \tag{9.1}$$

PROPOSITION 9.1. (a) *$R(x_1, x_2)$ is an upper semi-continuous function of the pair (x_1, x_2).*

(b) *$R(gx_1, gx_2) = R(x_1, x_2)$ if $g \in G$.*

(c) *For fixed $z \in Z$, the restriction of R to $X_z \times X_z$ defines a semi-metric on X_z.*

(d) *For every ϵ there is a δ such that whenever x_1, x_2, x'_1, x'_2 are in the same fibre and $d(x_1, x'_1) < \delta$, $d(x_2, x'_2) < \delta$ then $|R(x_1, x_2) - R(x'_1, x'_2)| < \epsilon$.*

Proof. For (a) we recall that if ϕ is an arbitrary function on a topological space and $\psi(x) = \limsup_{u \to x} \phi(u)$, then $\psi(x)$ is upper semi-continuous.

To prove (b) we notice that $S(x_1, x_2)$ is invariant under G. In particular $S(\tau x_1, \tau x_2) = S(x_1, x_2)$ for $\tau \in T$. Now if $(u_1, u_2) \to (x_1, x_2)$ then $(\tau u_1, \tau u_2) \to (\tau x_1, \tau x_2)$ by the continuity of τ, and from this it follows by (9.1) that $R(\tau x_1, \tau x_2) = R(x_1, x_2)$. Now for any g, (gx_1, gx_2) is a limit of a sequence $(\tau_n x_1, \tau_n x_2)$, so by (a), $R(gx_1, gx_2) \geqq R(x_1, x_2)$. Applying the same argument to g^{-1} we see that we must have $R(gx_1, gx_2) = R(x_1, x_2)$.

(c). By Theorem 7.7 (a), $\sigma(x_1, x_2)$ defines a semi-metric on X_0. It follows that $S(x_1, x_2)$ defines a semi-metric on each fibre. We wish to show that the same holds for $R(x_1, x_2)$. It is only necessary to establish the triangle inequality. So let $x_1, x_2, x_3 \in X_z$. We can find $(u_n, w_n) \to (x_1, x_3)$ with $S(u_n, w_n) \to R(x_1, x_3)$. Let $z_n = \pi(u_n) = \pi(w_n)$. Then $z_n \to \pi(x_2)$. By the corollary to Theorem 8.1, there is a sequence $\{v_n\}$ in X with $v_n \to x_2$ and $\pi(v_n) = z_n$. Then u_n, v_n, w_n are on one fibre and $S(u_n, w_n) \leqq S(u_n, v_n) + S(v_n, w_n)$. Letting $n \to \infty$ we find $R(x_1, x_3) \leqq R(x_1, x_2) + R(x_2, x_3)$ as was to be shown.

For the proof of (d), since $R(x_1, x_2)$ is a semi-metric, it suffices to show that for some δ if $d(x_1, x_2) < \delta$ where $\pi(x_1) = \pi(x_2)$, then $R(x_1, x_2) < \epsilon$. Let δ be chosen in accordance with Theorem 7.7 (d); that is, $F(u_1, u_2) < \delta$ for $u_1, u_2 \in X_0$ implies that $\sigma(u_1, u_2) < \epsilon$. Now if $d(x_1, x_2) < \delta$, then *a fortiori* $F(x_1, x_2) < \delta$. If $(u_1^{(n)}, u_2^{(n)}) \to (x_1, x_2)$ then $\limsup F(u_1^{(n)}, u_2^{(n)}) < \delta$. We may suppose that $R(x_1, x_2) = \lim S(u_1^{(n)}, u_2^{(n)}) = \lim \sigma(g_n u_1^{(n)}, g_n u_2^{(n)})$ where the g_n are chosen so that $g_n u_1^{(n)} \in X_0$. Since

$$\limsup F(g_n u_1^{(n)}, g_n u_2^{(n)}) = \limsup F(u_1^{(n)}, u_2^{(n)}) < \delta$$

it follows that $\lim \sigma(g_n u_1^{(n)}, g_n u_2^{(n)}) < \epsilon$. This completes the proof of the proposition.

We now strengthen (a) of the foregoing proposition.

PROPOSITION 9.2. *$R(x_1, x_2)$ is a continuous function on the subset of $X \times X$ defined by $\pi(x_1) = \pi(x_2)$.*

Proof. Let X^* denote the closed subset of $X \times X$ of pairs (x_1, x_2) with $\pi(x_1) = \pi(x_2)$. The flow (X^*, T) is defined as a component of the flow $(X \times X, T)$, and the Ellis group for this flow is G. (§ 3). Let $\pi^*(x_1, x_2) = \pi(x_1)$, so that $(Z, T) = \pi^*(X^*, T)$. The function $R(x_1, x_2)$ is an upper semi-continuous function on X^* and so it has points of continuity in every relatively open set of X^*. Let Λ_a denote the set of points at which R has discontinuity a:

$$(9.2) \qquad \Lambda_a = \{x^* \in X^* : R(x^*) - \liminf_{u^* \to x^*} R(u^*) \geqq a\}.$$

Λ_a is a closed set. Since $R(\tau x^*) = R(x^*)$ for $\tau \in T$, it follows that $T\Lambda_a \subset \Lambda_a$. Thus (Λ_a, T) is again a distal flow. Since (Z, T) is minimal, if Λ_a is non-empty we must have $Z = \pi^*(\Lambda_a)$, so that (Z, T) is a subflow of (Λ_a, T). Now if R is not continuous then some Λ_a is, in fact, non-empty.

Define a metric d^* in X^* by

$$d^*((x_1, x_2), (x'_1, x'_2)) = d(x_1, x'_1) + d(x_2, x'_2).$$

Also, for $\delta > 0$, set

$$\Lambda_a^\delta = \{x^* \in X^* : \exists x_1^* \in \Lambda_a \text{ with } \pi^*(x'_1) = \pi^*(x^*) \text{ and } d^*(x_1^*, x^*) < \delta\}.$$

Clearly $\Lambda_a \subset \Lambda_a^\delta$. We shall show that Λ_a^δ contain a neighborhood of each point $x_0^* \in \Lambda_a$. Suppose then that $x_n^* \to x_0^*$; then $\pi^*(x_n^*) \to \pi^*(x_0^*)$. Now because (Z, T) is minimal, and (Λ_a, T) is distal, we may apply Theorem 8.1 and its corollary to the map π^*. We conclude that there exists a sequence $\{u_n^*\}$, with $u_n^* \to x_0^*$, $\pi^*(u_n^*) = \pi^*(x_n^*)$ and $u_n^* \in \Lambda_a$. Since $(u_n^*, x_n^*) \to (x_0^*, x_0^*)$, $d^*(u_n^*, x_n^*) \to 0$, which shows that ultimately $x_n^* \in \Lambda_a^\delta$.

By Proposition 9.1 (d), there exists a δ such that if $\pi^*(u^*) = \pi^*(v^*)$ and $d^*(u^*, v^*) < \delta$, then $|R(u^*) - R(v^*)| < a/4$. Form Λ_a^δ for this δ. Λ_a^δ contains an open set of X^* and so R will have a point u^* of continuity in Λ_a^δ. Then there is an $x^* \in \Lambda_a$ with $\pi^*(x^*) = \pi^*(u^*)$ and $d^*(u^*, x^*) < \delta$. Since $x^* \in \Lambda_a$, there is a sequence $x_n^* \to x^*$ with $\lim R(x_n^*) \leqq R(x^*) - a$. Now apply the corollary to Theorem 8.1 to the map $\pi^* : X^* \to Z$. Since $\pi^*(x_n^*) \to \pi^*(x^*) = \pi^*(u^*)$, we can find $u_n^* \to u^*$ with $\pi^*(u_n^*) = \pi^*(x_n^*)$. Then $(u_n^*, x_n^*) \to (u^*, x^*)$ and since $d^*(u^*, x^*) < \delta$, ultimately $d^*(u_n^*, x_n^*) < \delta$, and also $R(u_n^*) < R(x_n^*) + a/4 < R(x^*) - 3a/4 < R(u^*) - a/2$. Since $u_n^* \to u^*$ and R is continuous at u^*, this is a contradiction. This proves that Λ_a must be empty and R is everywhere continuous.

PROPOSITION 10.3. *If (X, T) is a non-trivial extension of (Z, T), then the function $R(x_1, x_2)$ does not reduce to a constant.*

Proof. If $R(x_1, x_2)$ is constant it vanishes everywhere. But on X_0, $R(x_1, x_2) \geqq S(x_1, x_2) = \sigma(x_1, x_2)$, and this will not vanish everywhere unless M reduces to a single pt. The proposition now follows from Theorem 7.6.

10. Construction of (Y, T). We have $\pi(X,T) = (Z,T)$ and for pairs of points $x_1, x_2 \in X$ with $\pi(x_1) = \pi(x_2)$, there is defined a function $R(x_1, x_2)$ satisfying the following:

(10.1) (a) R is a continuous function.

(b) On each fibre of X with respect to Z, R defines a semi-metric.

(c) For each $g \in G$, $R(gx_1, gx_2) = R(x_1, x_2)$.

(d) R is not constant.

Since R is a semi-metric it is natural to introduce an equivalence relation: $x_1 \sim x_2$ if $\pi(x_1) = \pi(x_2)$ and $R(x_1, x_2) = 0$. It is clear that this is an equivalence relation. Moreover by (c) we shall have $x_1 \sim x_2$ implies $gx_1 \sim gx_2$. Let Y denote the set of equivalence classes for this relation, and let π_1 be the natural map $\pi_1 \colon X \to Y$ taking each point of X into its equivalence class. Since x_1 and x_2 are equivalent only if $\pi(x_1) = \pi(x_2)$, we may define a map $\pi_2 \colon Y \to Z$ by

(10.2) $$\pi_2(\pi_1(x)) = \pi(x).$$

We then have

(10.3) $$X \xrightarrow{\pi_1} Y \xrightarrow{\pi_2} Z$$

and $\pi_2 \circ \pi_1 = \pi$.

We endow Y with a topology by designating $U \subset Y$ as open if and only if $\pi_1^{-1}(U)$ is open in X. Then π_1 is also continuous. π_2 is also continuous, since $\pi_1^{-1}(\pi_2^{-1}(W)) = \pi^{-1}(W)$, so if W is open in Z, $\pi_2^{-1}(W)$ is open in Y.

PROPOSITION 10.1. *Y is a compact, Hausdorff space.*

Proof. Since π_1 is continuous, Y is compact. To prove it is Hausdorff, suppose $y_1, y_2 \in Y$ and $y_1 \neq y_2$. Suppose first that $\pi_2(y_1) \neq \pi_2(y_2)$. Then since Z is a Hausdorff space we can separate $\pi_2(y_1)$ and $\pi_2(y_2)$ by open sets W_1 and W_2 with $W_1 \cap W_2 = \emptyset$, and since π_2 is continuous $\pi_2^{-1}(W_1)$ and $\pi_2^{-1}(W_2)$ are again open nonoverlapping sets in Y which separate y_1 and y_2. So we may confine ourselves to the case $\pi_2(y_1) = \pi_2(y_2)$. Let $\Delta_1 = \pi_1^{-1}(y_1)$, $\Delta_2 = \pi_1^{-1}(y_2)$. Since $y_1 \neq y_2$ there is a constant $a > 0$ such that $R(x_1, x_2) = a$ for any $x_1 \in \Delta_1$ and $x_2 \in \Delta_2$. Since $R(x_1, x_2)$ is continuous we may enclose Δ_1 in an open set U_1, and Δ_2 in U_2 such that $R(u_1, u_2) > a/2$ whenever $u_1 \in U_1$, $u_2 \in U_2$. Now let V_1 be the set of points not equivalent $(\sim)$ to any point outside of U_1, and define V_2 analogously. Clearly $\Delta_1 \subset V_1 \subset U_1$, $\Delta_2 \subset V_2 \subset U_2$. We claim that V_1 and V_2 are open. For if $v_n \notin V_1$ and $v_n \to v$ then $v_n \sim u_n \notin U_1$. Passing to a subsequence we may suppose that $u_n \to u$. $u \in U_1$ since U_1 is open. Then $R(u, v) = \lim R(u_n, v_n) = 0$ so that $u \sim v$ and $v \notin V_1$. This shows that

V_1 is open and similarly V_2 is open. Also any point equivalent to a point of V_1 is clearly in V_1, so $\pi_1^{-1}(\pi_1(V_1)) = V_1$; similarly $\pi_2^{-1}(\pi_2(V_2)) = V_2$. Hence $\pi_1(V_1)$ and $\pi_1(V_2)$ are open in Y, $\pi_1(V_1) \cap \pi_1(V_2) = \emptyset$ (since $V_1 \cap V_2 = \emptyset$) and $y_1 \in \pi_1(V_1)$, $y_2 \in \pi_1(V_2)$. This completes the proof.

Let X_0 be a fibre of X over Z. Let $Y_0 = \pi_1(X_0)$ be its image in Y. In Y_0 we define a metric by

$$\rho(\pi_1(x_1), \pi_1(x_2)) = R(x_1, x_2) \qquad x_1, x_2 \in X_0. \tag{10.4}$$

This is unambiguous since $R(x_1, x'_1) = 0$ and $R(x_2, x') = 0$ implies $R(x_1, x_2) = R(x'_1, x'_2)$. Since R defined a semi-metric in X_0 it follows that ρ defines a metric in Y_0. (The continuity of ρ is a consequence of the continuity of R.)

Now let $G_0 = \{g : gX_0 \subset X_0\}$. For each $g \in G$ we may define $w(g)$ on Y by

$$w(g)\pi_1(x) = \pi_1(gx) \tag{10.5}$$

and this is unambiguous since $x_1 \sim x_2$ implies $gx_1 \sim gx_2$. Let $H_0 = w(G_0)$. Then H_0 takes Y_0 into itself. We claim that H_0 acts on Y_0 by isometries. For

$$\begin{aligned}\rho(w(g)\pi_1(x_1), w(g)\pi_1(x_2)) &= \rho(\pi_1(gx_1), \pi_1(gx_2)) \\ &= R(gx_1, gx_2) = R(x_1, x_2) = \rho(\pi_1(x_1), \pi_1(x_2)).\end{aligned}$$

Moreover since G_0 is transitive on X_0, H_0 will be transitive on Y_0. Thus Y_0 is a homogeneous, compact metric space.

Since T acts on X continuously, it may be seen that T acts on Y continuously. For let $\tau_n \to \tau$, $y_n \to y$, and $w(\tau_n)y_n \to y'$. Let $y_n = \pi_1(x_n)$. Passing to a subsequence we may suppose that $x_n \to x$ in X. Then $y = \pi_1(x)$.

$$w(\tau_n)y_n = w(\tau_n)\pi_1(x_n) = \pi_1(\tau_n x_n) \to \pi_1(\tau x) = w(\tau)\pi_1(x) = w(\tau)y,$$

or $y' = w(\tau)y$. This shows that whenever $\tau_n \to \tau$, $y_n \to y$, then $w(\tau_n)y_n \to w(\tau)y$. If we now define the action of T on Y by $\tau y = w(\tau)y$, then it follows that (Y, T) is a flow.

THEOREM 10.2. *Y is a Y_0-bundle over Z and (Y, T) is an isometric extension of (Z, T).*

Proof. To show that Y is a Y_0-bundle over Z, it is necessary to show the existence of a function $\rho(y_1, y_2)$ satisfying the conditions of Def. 2.1. We define ρ by (10.4). It is then continuous and defines a metric on each fibre of Y over Z. Since $R(gx_1, gx_2) = R(x_1, x_2)$, it follows that ρ is invariant under $w(G)$. $w(G)$ is transitive on Y since G is transitive on X ((X, T) is minimal), and this shows that each fibre is isometric to Y_0. This shows that Y is a Y_0-bundle over Z.

We have $(Z,T)=\pi_2(Y,T)$. To show that (Y,T) is an isometric extension of (Y,T) it is necessary to show that ρ is invariant under T. But ρ is invariant under $w(G)$ and the operations of T correspond to $w(T)\subset w(G)$. This completes the proof of the theorem.

Let us observe that under (10.1 d), (Y,T) is a non-trivial extension of (Z,T). For Y_0 cannot reduce to a single point unless R vanishes identically. We have thereby proved

THEOREM 10.3. *If (X,T) is a minimal distal flow, and (Z,T) is a proper subflow, then there exists a subflow (Y,T) of (X,T) which is a non-trivial isometric extension of (Z,T).*

Combining this with Theorem 4.2 we have proved

THEOREM 2.4. *Every minimal distal flow is quasi-isometric.*

11. The Abelian case and eigenfunctions of a flow. When the group T is abelian then there is an alternative approach to the proof of Theorem 2.4 which we shall indicate in this section. This approach does not actually yield a complete proof of Theorem 2.4 even in this case. Nevertheless it is of some interest as far as it goes.

If (X,T) is a quasi-isometric flow, then it is obtained by a succession (possibly transfinite) of isometric extensions from the trivial flow. Let (X_1,T) denote the first such extension (which is non-trivial), so that $(X_1,T)=\pi_1(X,T)$ is an isometric extension of the trivial flow. Clearly this means that T acts on X_1 by isometries. The Ellis group of (X_1,T) is then a compact group of isometries of X_1. If T is abelian then this group is abelian and X_1 itself is a compact abelian group and T acts by rotations. Let ϕ denote a character on the group X_1. Imbedding T in X_1 we have $\phi(\tau x)=\phi(\tau)\phi(\mathrm{x})$ where $\tau\in T$, $x\in X_1$. Thus ϕ is an eigenfunction of every transformation of T. If we set $\Phi(x)=\phi(\pi_1(x))$ we will have $\Phi(\tau x)=\phi(\tau)\Phi(x)$. Thus if (X,T) is a q.i. flow and T is abelian, there exist non-constant continuous eigenfunctions for the flow. It is possible to show conversely that if a flow (X,T) possesses non-constant continuous eigenfunctions, then it has a non-trivial subflow (X_1,T) which is equicontinuous, i.e. an isometric extension of the trivial flow.

We will now show directly (i.e. without invoking Theorem 2.4) that if (X,T) is a minimal distal flow, it possesses non-constant continuous eigenfunctions, if T is abelian.

Let M be the set of non-negative regular Borel measures on X satisfying $\mu(X)=1$. M is a convex set which is compact with respect to the weak* topology of M as a subset of the dual space to $C(X)$, the continuous functions on X. Define $\tau\mu$ by

$$\int f(x)\,d\tau\mu(x)=\int f(\tau x)\,d\mu(x).$$

This is an affine transformation of $M \to M$. By the Markov-Kakutani fixed point theorem ([3]) there exists some measure μ satisfying $\tau\mu = \mu$ for all $\tau \in T$, if T is abelian. μ will henceforth denote an invariant measure. Form the Hilbert space $L^2(X, \mu)$ and define the operator A by

$$Af(x) = \int F(x, y) f(y)\, d\mu(y) \tag{11.1}$$

where F is the function defined in § 5.

Since $F(x, y) = F(y, x)$ and $F(x, y)$ is bounded, A is a self adjoint Hilbert-Schmidt operator. For each $\tau \in T$ define the operator L_τ by $L_\tau f(x) = f(\tau x)$. We have then

$$\begin{aligned} AL_\tau f(x) &= \int F(x, y) f(\tau y)\, d\mu(y) \\ &= \int F(\tau x, \tau y) f(\tau y)\, d\mu(y) = \int F(\tau x, y) f(y)\, d\tau\mu(y) \\ &= \int F(\tau x, y) f(y)\, d\mu(y) = L_\tau A f(x). \end{aligned}$$

So each L_τ commutes with A. In addition each L_τ is a unitary operator on $L^2(X, \mu)$. Now by the theory of completely continuous operators ([8]), $L^2(X, \mu)$ decomposes with a direct sum

$$L^2(X, \mu) = H_0 \oplus H_{\lambda_1} \oplus H_{\lambda_2} \oplus \cdots \oplus H_{\lambda_n} \oplus \cdots$$

where H_λ denotes the set of elements $f \in L^2(X, \mu)$ for which $Af = \lambda f$. Each of the spaces H_λ, for $\lambda \neq 0$, is finite dimensional. Since the L_τ commutes with A, each L_τ takes H_λ into H_λ, and so T maps into a commutative group of unitary operators on a finite dimensional space H_λ $(\lambda \neq 0)$. This means that H_λ has a basis consisting of eigenfunctions for all the L_τ. Now it is also known that the kernel $F(x, y)$ of A can be approximated in $L^2(X \times X, \mu \times \mu)$ by sums of products of functions in the various H_λ with $\lambda \neq 0$. So unless F is almost everywhere (measure theoretically) a constant, the H_λ for some $\neq 0$ will have positive dimensions and the L_τ will possess non-constant eigenfunctions. These will be measurable eigenfunctions of the flow. Now if $F(x, y)$ is constant almost everywhere, since $F(x, y) \leqq d(x, y) < \epsilon$ on a set of positive measure, if $\epsilon > 0$, it follows that $F(x, y) = 0$ a.e. But then, unless X is trivial, there must be pairs of distinct points x and y with $F(x, y) = 0$, contrary to supposition. This proves the existence of measurable eigenfunctions.

Now if $\phi \in H_\lambda$, $\lambda \neq 0$ then it has a representation as a function in the first Baire class:

$$\phi(x) = \lambda^{-1} \int F(x, y) \phi(y)\, d\mu(y)$$

since F is upper semi-continuous. A result of Kakutani (unpublished) implies that if ϕ is an eigenfunction of a minimal flow then ϕ is continuous. In fact

if ϕ had a discontinuity at x, since $\phi(\tau x) = \theta(\tau)\phi(x)$, $|\theta(\tau)| = 1$, it follows that it has a discontinuity of the same magnitude at τx. Hence it has a discontinuity of this magnitude on a dense set, and therefore ϕ has no points of continuity. Since this cannot happen if ϕ belongs to the first Baire class, ϕ must be continuous everywhere. This establishes our assertion.

In case (X, T) is not just minimal, distal, but also strictly ergodic (i. e. the invariant measure μ is unique), this method may be extended to show that (X, T) is a quasi-isometric flow. This is done by showing that each subflow (Z, T) of (X, T) possesses a non-trivial isometric extension in (X, T). This extension is obtained by studying operators A_z defined on each fibre of X over Z in a manner similar to (11.1). This method could be used in the non-abelian case if we knew that whenever (X, T) is distal, there is some invariant measure μ for T on X. As a matter of fact, this is the case, as we will show in the next section. However the proof depends upon Theorem 2.4, so this does not yield an alternative proof of that theorem. In any case, this procedure has only been successful if we add the hypothesis of strict ergodicity.

An interesting consequence of the existence of eigenfunctions on X when (X, T) is minimal distal is the following:

THEOREM 11.1. *If X is simply connected, then X does not admit of a minimal distal flow for any locally compact abelian group T.*

Proof. If (X, T) is minimal distal, there exists a non-constant continuous map Φ of X into the unit circle with $\Phi(\tau x) = \phi(\tau)\Phi(x)$ for $x \in X$, $\tau \in T$. If X is simply connected, we may write $\Phi(x) = e^{i\psi(x)}$ where $\psi(x)$ is real valued and continuous on X. Then

$$e^{i[\psi(\tau x)-\psi(x)]} = \phi(\tau).$$

Since $\psi(\tau x) - \psi(x)$ is continuous as a function of x, it must be constant for each τ. Since ψ is bounded, this implies $\psi(\tau x) = \psi(x)$ for each x and τ. But then ψ is constant and so also is Φ. This proves the theorem.

In particular, the n-sphere, for $n > 1$, cannot support a minimal distal flow.

12. Existence of an invariant measure for a distal flow. In this section we shall show that if T is a locally compact group of homeomorphisms of a compact metric space such that (X, T) is a distal flow, then there exists on X a non-negative Borel measure which is left fixed by T.

It is necessary to elaborate on some of the notions referred to in the preceding section. Let $C(X)$ denote the Banach space of continuous functions

on X with the sup norm. The dual space $C^*(X)$ may be identified with the set of Borel measures on X. Namely if μ is a (Borel) measure on X we set

$$(12.1) \qquad \mu(f) = \int_X f(x)\, d\mu(x).$$

Now let τ be homeomorphism of X. For every $f \in C(X)$ we set $f^\tau(x) = f(\tau x)$. If $\mu \in C^*(X)$ we define $\tau\mu$ by

$$(12.2) \qquad \tau\mu(f) = \mu(f^\tau).$$

We say μ is invariant (under T) if $\tau\mu = \mu$ for all $\tau \in T$. In the space of measures $C^*(X)$ we shall not be concerned with the norm topology, but rather with the weak* topology. That is, we say $\mu_n \to \mu$ if, for each $f \in C(X)$, $\mu_n(f) \to \mu(f)$. We shall say that $\mu \in C^*(X)$ is a *probability measure* on X if $\mu \geqq 0$ and $\mu(X) = 1$.

LEMMA 12.1. *Let M be a homogeneous compact metric space. There is a unique probability measure μ on M satisfying $\gamma\mu = \mu$ for every isometry γ of M.*

Proof. Let Γ be the compact group of isometries of M and let m denote the normalized Haar measure on Γ. For each $x \in M$ define $m_x \in C^*(M)$ by $m_x(f) = \int_\Gamma f(\gamma x)\, dm(\gamma)$. Since m is right invariant, $m_{\gamma' x} = m_x$ for $\gamma' \in \Gamma$. Since Γ is transitive, $m_x = m_y$ for any $x, y \in M$. Since m is left invariant, m_x is invariant under Γ. Denote the common measure m_x by μ. Now suppose that μ' is any probability measure invariant under Γ. Then

$$\mu'(f) = \int_M f(z)\, d\mu'(z) = \int_M f(z)\, d\gamma\mu'(z) = \int_\Gamma \int_M f(z)\, d\gamma\mu'(z)\, dm(\gamma)$$
$$= \int_\Gamma \int_M f(\gamma z)\, d\mu'(z)\, dm(\gamma) = \int_M m_z(f)\, d\mu'(z) = \mu(f).$$

This proves the lemma.

PROPOSITION 12.1. *Suppose (Y, T) is a flow such that T leaves invariant a probability measure on Y. If (X, T) is an isometric extension of (Y, T), then T leaves invariant a probability measure on X.*

Proof. Let X be an M-bundle over Y and suppose $(Y, T) = \pi(X, T)$. Each fibre X_y of X over Y is isometric to M and hence possesses a unique probability measure invariant under its isometries. Denote this measure by μ_y. Each μ_y is a measure on $X \supset X_y$. Now let $\tau \in T$ and consider the measure $\tau\mu_y$. Since $\tau X_y \subset X_{\tau y}$, $\tau\mu_y$ is concentrated on $X_{\tau y}$. Suppose that γ is an

isometry of $X_{\tau y}$. By Definition 2.2, it follows that τ is an isometry of X_y onto $X_{\tau y}$. Hence $\tau^{-1}\gamma\tau$ is an isometry of X_y and so $\tau^{-1}\gamma\tau\mu_y = \mu_y$, or $\gamma\tau\mu_y = \tau\mu_y$. So $\tau\mu_y$ is a probability measure on $X_{\tau y}$ invariant under every isometry of $X_{\tau y}$. Hence $\tau\mu_y = \mu_{\tau y}$.

Next we shall show that μ_y is a continuous function from Y to $C^*(X)$ relative to the weak* topology of the latter. Namely suppose $y_n \to y$. For each n let $\gamma_n : X_y \to X_{y_n}$ be an isometry. As maps of $X_y \to X$ the γ_n are equicontinuous and so, passing to a subsequence and renumbering, we may suppose that $\gamma_n \to \gamma$ uniformly. γ will take X_y into itself. Let ρ define the metric on the fibres of X with respect to Y. Then $\rho(\gamma_n x_1, \gamma_n x_2) = \rho(x_1, x_2)$, and since ρ is continuous, $\rho(\gamma x_1, \gamma x_2) = \rho(x_1, x_2)$, and we see that γ is an isometry of X_y. Passing to a further subsequence of $\{y_n\}$ we may suppose that μ_{y_n} converges, say to μ'. Let μ' be an isometry of X_y. Then $\gamma_n\gamma'\gamma_n^{-1}$ is an isometry of X_{y_n} so that $\gamma_n\gamma'\gamma_n^{-1}\mu_{y_n} = \mu_{y_n}$. Since $\gamma_n \to \gamma$ uniformly, $\gamma_n\gamma'\gamma_n^{-1} \to \gamma\gamma'\gamma^{-1}$ uniformly and since $\mu_{y_n} \to \mu'$, $\gamma_n\gamma'\gamma_n^{-1}\mu_{y_n} \to \gamma\gamma'\gamma^{-1}\mu'$. Thus $\gamma\gamma'\gamma^{-1}\mu' = \mu'$ or $\gamma'\gamma^{-1}\mu' = \gamma^{-1}\mu'$, and $\gamma^{-1}\mu'$ is invariant under any isometry of X_y. Hence $\mu' = \gamma^{-1}\mu'$ and μ' must be identical with μ_y. What we have shown is that if $y_n \to y$ and μ_{y_n} converges, the limit must be μ_y. Since the probability measures on X form a compact set, it follows that μ_y is a continuous map from Y to $C^*(X)$.

Now let ν denote an invariant probability measure on Y. We define a measure on X by setting

$$\nu(\phi) = \int \mu_y(\phi)\, d\nu(y). \tag{12.3}$$

The integrand is a continuous function of y and so the integral in (12.3) is well-defined. It is readily verified that this defines a positive linear functional on $C(X)$ with $\bar{\nu}(1) = 1$, so that ν is a probability measure. We claim that ν is invariant under T. Let $\tau \in T$:

$$\begin{aligned} \tau\bar{\nu}(\phi) &= \bar{\nu}(\phi^\tau) = \int \mu_y(\phi^\tau)\, d\nu(y) = \int \tau\mu_y(\phi)\, d\nu(y) \\ &= \int \mu_{\tau y}(\phi)\, d\nu(y) = \int \mu_y(\phi)\, d^\tau\nu(g) = \int \mu_y(\phi)\, d\nu(y) \\ &= \bar{\nu}(\phi). \end{aligned}$$

This completes the proof of the proposition.

Suppose π is a continuous map of X into Y. For any $f \in C(Y)$ set $f^\pi(x) = f(\pi x)$. Then if μ is a measure on X let $\pi\mu$ be defined by $\pi\mu(f) = \mu(f^\pi)$.

LEMMA 12.2. *Let (X, T) be a flow and μ a measure on X invariant under T. If $(Y, T) = \pi(X, T)$ is a subflow, then $\pi\mu$ is a measure on Y invariant under T.*

Proof. First we remark that for $\tau \in T$, $(f^\tau)^\pi = (f^\pi)^\tau$, since for $x \in X$, $\tau\pi(x) = \pi(\tau x)$. It follows that $\tau\pi\mu = \pi\tau\mu$. But $\tau\mu = \mu$ so $\pi\tau\mu = \pi\mu$ and $\tau\pi\mu = \pi\mu$. This proves the lemma.

PROPOSITION 12.2. *Let Σ be a totally ordered family of subflows $\{(X_\alpha, T)\}$ of (X,T) such that (X,T) is a limit of the flows of Σ (Def. 2.3). If for each α, X_α carries a probability measure invariant under T, then X does so as well.*

Proof. Let A be the index set for Σ: $\Sigma = \{(X_\alpha, T), \alpha \in A\}$ so that A is a totally ordered set, and if $\alpha, \beta \in A$, $\alpha > \beta$, then $(X_\beta, T) \prec (X_\alpha, T)$. Suppose $(X_\alpha, T) = \pi_\alpha(X,T)$. If $\alpha > \beta$; there is a map $\pi_\beta{}^\alpha \colon X_\alpha \to X_\beta$ with $\pi_\beta = \pi_\beta{}^\alpha \circ \pi_\alpha$.

Let ν_α be a probability map on X_α invariant under T. $\nu_\alpha \in C^*(X_\alpha)$, and the map $f \to f^{\pi_\alpha}$ imbeds $C(X_\alpha)$ in $C(X)$. So ν_α is a linear functional of norm 1 defined on a subspace of $C(X)$. It extends to some linear functional on all of $C(X)$ without increasing its norm. Let $\bar\nu_\alpha$ denote this extension. Then $\bar\nu_\alpha(1) = 1$ and $\| \bar\nu_\alpha \| = 1$ which implies that $\bar\nu_\alpha$ is a probability measure. We have $\pi_\alpha \bar\nu_\alpha = \bar\nu_\alpha$. For $\beta < \alpha$, $\pi_\beta \nu_\alpha = \pi_\beta{}^\alpha \pi_\alpha \nu_\alpha = \pi_\beta{}^\alpha \nu_\alpha$. Since ν_α is invariant for (X_α, T) and $(X_\beta, T) = \pi_\beta{}^\alpha(X_\alpha, T)$, by Lemma 12.2, $\pi_\beta \nu_\alpha$ is invariant for (X_β, T). As a result, for each α we have a probability measure $\bar\nu_\alpha$ on X with $\pi_\beta(\nu_\alpha)$ an invariant measure on X_β for all $\beta > \alpha$. Let P_α denote the set of all probability measures on X with this property. We see easily that P_α is a compact set. Clearly $P^\alpha \subset P_{\alpha'}$ if $\alpha > \alpha'$. It follows that there is a measure $\mu \in \cap P_\alpha$.

We now claim that this measure μ is invariant under T. To see this, let B_α denote the algebra of functions on X of the form f^{π_α} where $f \in C(X_\alpha)$. Since A is totally ordered, $\cup B_\alpha$ is an algebra. Since (X,T) is a limit of $\{(X_\alpha, T)\}$, it follows that $\cup B_\alpha$ separates points in X. It is therefore dense in $C(X)$. For $\phi \in B_\alpha$, $\phi = f^{\pi_\alpha}$, $\mu(\phi) = \pi_\alpha\mu(f)$. Since $\pi_\alpha\mu$ is invariant under T, it follows that if $\tau \in T$, $\tau\mu(\phi) = \mu(\phi)$. Since this is the case for a dense subset of $C(X)$, we conclude that μ is invariant. This proves the proposition.

Applying Propositions 12.1 and 12.2 and using transfinite induction we conclude that for any quasi-isometric flow there exists an invariant probability measure. By Theorem 2.4, this will be true for any minimal distal flow. Since any flow has minimal components, the result is true for any distal flow.

THEOREM 12.3. *If (X,T) is a distal flow where T is arbitrary locally compact group, then there exists a probability measure on X, invariant under T.*

13. Order of a distal flow. In the definition of a quasi-isometric flow, there occurred an ordinal η corresponding to the "number" of successive isometric extension needed to pass from the trivial flow to the given one. Clearly this ordinal is not unique. If X is the circle $\{|\zeta| = 1\}$ and T is a

group of rotations of X, then $\pi\colon X \to X'$ defined by $\pi(\zeta) = \zeta^2$ defines a subflow (X',T) of (X,T). (X',T) is an isometric extension of the trivial flow and (X,T) is an isometric extension of (X',T), so that η could be taken as 1 or 2 (or any finite ordinal) for (X,T). Nevertheless, the least ordinal which could occur is well defined.

Definition 13.1. If (X,T) is a q.i. flow, let η be the smallest ordinal such that we can write $(X,T) = (X_\eta, T)$ where $\{(X_\xi, T), \xi \leqq \eta\}$ is a quasi-isometric system (§ 4). η is referred to as the *order* of (X,T).

PROPOSITION 13.1. *Let (X,T) be a distal flow and $(Z,T) = \pi(X,T)$ a subflow. There exists an isometric extension (Y,T) of (Z,T) which is a subflow of (X,T) such that if (Y',T) is any flow with these two properties, then (Y',T) is a subflow of (Y,T).*

Proof. Let $\{(Y_\alpha, T)\}$ range over the set of all flows which are subflows of (X,T) and isometric extensions of (Z,T). Set $(Y_\alpha, T) = \pi_\alpha(X,T)$. For each α, there is a function $\rho_\alpha(y_1, y_2)$ defined on the fibres of Y_α with respect to Z, and satisfying the conditions of Definitions 2.1 and 2.2. Set $R_\alpha(x_1, x_2) = \rho_\alpha(\pi_\alpha x_1, \pi_\alpha x_2)$ when $\pi x_1 = \pi x_2$. $R_\alpha(x_1, x_2)$ satisfies:

(a) R_α is a continuous function.
(b) $R_\alpha(x_1, x_2)$ defines a semi-metric on each fibre of X over Z
(c) For $g \in G$, the Ellis group of (X,T), $R_\alpha(gx_1, gx_2) = R_\alpha(x_1, x_2)$

If $X^* = \{(x_1, x_2) : \pi(x_1) = \pi(x_2)\}$ then X^* is a compact metric space. In particular $C(X^*)$ is separable, and it follows that there exists a countable subset of $\{R_\alpha\}$ which is dense in the entire set. Let $\{R_{\alpha_n}\}$ denote this countable subset. Each R_{α_n} is bounded, so for some sequence of positive terms $\{c_n\}$, $\Sigma c_n R_{\alpha_n}$ converges uniformly. We set

$$R(x_1, x_2) = \Sigma\, c_n R_{\alpha_n}(x_1, x_2). \tag{13.1}$$

R will again satisfy (a), (b), and (c) above. We now use the results of § 10. We define $x_1 \sim x_2$ if $\pi(x_1) = \pi(x_2)$ and $R(x_1, x_2) = 0$, and let Y denote the set of equivalence classes. By what we have shown in § 10, we may define a flow (Y,T) which is a subflow of (X,T) and an isometric extension of (Z,T). We claim that (Y,T) extends each (Y_α, T). To show this it is necessary to show that if $x_1 \sim x_2$, then $\pi_\alpha(x_1) = \pi_\alpha(x_2)$ for each α. But if $x_1 \sim x_2$, then $R(x_1, x_2) = 0$ and by (13.1), $R_{\alpha_n}(x_1, x_2) = 0$. Since $\{R_{\alpha_n}\}$ is dense in $\{R_\alpha\}$, we find $R_\alpha(x_1, x_2) = 0$ for each α. Since ρ_α is a metric on the fibres of Y_α, this can only happen if $\pi_\alpha x_1 = \pi_\alpha x_2$. This proves the proposition.

Definition 13.2. Let (X,T) be a distal flow and (Z,T) a subflow. The

flow (Y,T) of Prop. 13.1 will be called the *maximal isometric extension of* (Z,T) in (X,T).

(Y,T) is clearly unique up to isomorphism, so this definition is unambiguous.

Definition 13.3. A quasi-isometric system $\{(X_\xi,T),\xi\leqq\eta\}$ is said to be *normal if* each $(X_{\xi+1},T)$ is the maximal isometric extension of (X_ξ,T) in (X_η,T) where $\xi<\eta$.

LEMMA 13.1. *Let 4 flows be given as in the accompanying diagram where arrows indicate extensions. Suppose that* $w_1\circ\pi_1=w_2\circ\pi_2$ *and that whenever* y *and* y' *are distinct points of* Y, *either* $\pi_1(y)\neq\pi_1(y')$ *or* $\pi_2(y)\neq\pi_2(y')$. *Then if* (Z_1,T) *is an isometric extension of* (W,T), (Y,T) *will be an isometric extension of* (Z_2,T).

$$\begin{tikzcd}[column sep=small] & (Y,T) \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ (Z_1,T) \arrow[dr, "w_1"'] & & (Z_2,T) \arrow[dl, "w_2"] \\ & (W,T) & \end{tikzcd}$$

Proof. Let $\rho(z_1,z'_1)$ be the metric on the fibres of Z_1 over W. Set $\sigma(y,y')=\rho(\pi_1(y),\pi_1(y'))$ whenever $\pi_2(y)=\pi_2(y')$. Since we then have $w_2\circ\pi_2(y)=w_2\circ\pi_2(y')$ and $w_1(\pi_1(y)=w_1(\pi_1(y'))$, we see that σ is well defined. σ is clearly a semi-metric on the fibres of Y over Z_2, and is invariant under T. To show it is a metric we remark that if $\sigma(y_1,y')=0$ for $\pi_2(y)=\pi_2(y')$, then we must also have $\pi_1(y)=\pi_1(y')$, or $y=y'$. This completes the proof.

THEOREM 13.2. *If* $\{(X_\xi,T),\xi\leqq\eta\}$ *is a normal quasi-isometric system, then* (X_η,T) *has order* η.

Proof. Suppose $\{(X'_\xi,T),\xi\leqq\eta'\}$ is another q.i. system terminating with (X_η,T). We shall show that for each $\xi\leqq\eta$, (X'_ξ,T) is a subflow of (X_ξ,T). This is done by transfinite induction. Suppose $(X'_\xi,T)=\theta(X_\xi,T)$. Set $(X_\xi,T)=\pi_\xi(X_\eta,T)$, $(X'_\xi,T)=\pi'_\xi(X_\eta,T)$. Let Y be the subset of $X\times X'_{\xi+1}$ of points of the form $(\pi_\xi(x),\pi'_{\xi+1}(x))$ where $x\in X$. There is a flow (Y,T) which is then a subflow of (X,T) and we have $\pi(Y,T)=(X_\xi,T)$, $\pi'(Y,T)=(X'_{\xi+1},T)$. We then have the accompanying diagram. Since all

$$\begin{tikzcd}[column sep=small] & (Y,T) \arrow[dl, "\pi"'] \arrow[dr, "\pi'"] & \\ (X_\xi,T) \arrow[dr, "\theta"'] & & (X'_{\xi+1},T) \arrow[dl, "\theta'"] \\ & (X'_\xi,T) & \end{tikzcd}$$

the flows are the induced maps we have $\theta\circ\pi=\theta'\circ\pi$. Also, by the definition of (Y,T) it is clear that π and π' separate points in Y. We are therefore in a

position to use the foregoing lemma. Since θ' defines an isometric extension, π does so as well. But then (Y, T) is a subflow of $(X_{\xi+1}, T)$, the maximal isometric extension of (X_ξ, T) in (X_η, T). It follows that $(X'_{\xi+1}, T)$ is a subflow of $(X_{\xi+1}, T)$. Let Ω be the set of ordinals ξ for which $(X'_\xi, T) \prec (X_\xi, T)$. We have proven that $\xi \in \Omega$ implies $\xi + 1 \in \Omega$. If ξ is a limit ordinal, then (X'_ξ, T) is the limit of $\{(X'_\zeta, T), \zeta < \xi\}$ and similarly for (X_ξ, T). It follows that if $\zeta \in \Omega$ for $\zeta < \xi$ then $\xi \in \Omega$. Hence Ω contains all ordinals $\leqq \eta$. So $(X'_\eta, T) \prec (X_\eta, T)$. But $(X'_{\eta'}, T) = (X_\eta, T)$ so $\eta' \geqq \eta$ and this proves the theorem.

To illustrate this result, let us take X to be the infinite torus, $X = \{x = (\zeta_1, \cdots, \zeta_n, \cdots), |\zeta_i| = 1\}$, and $X_n = \pi_n(X)$, where π_n assigns to each x, its first n components. Let T be the integer group and let (X, T) be generated by

$$\tau_1(\zeta_1, \zeta_2, \cdots, \zeta_n, \cdots) = (e^{i\alpha}\zeta_1, \zeta_1\zeta_2, \cdots, \zeta_{n-1}\zeta_n, \cdots). \tag{13.2}$$

and (X_n, T) by

$$\tau_1(\zeta_1, \zeta_2, \cdots, \zeta_n, \cdots) = (e^{i\alpha}\zeta_1, \zeta_1\zeta_2, \cdots, \zeta_{n-1}\zeta_n). \tag{13.3}$$

Here we take α to be an irrational multiple of π. Clearly $(X_n, T) = \pi_n(X, T)$. Now consider the quasi-isometric system consisting of the (X_n, T) together with (X, T). If we can show that this is normal, it will follow that the order of (X, T) is the ordinal ω.

LEMMA 13.2. *The flow (X_n, T) is minimal.*

Proof. See [6, Theorem 2.1].

LEMMA 13.3. *The flow (X, T) is minimal.*

Proof. This is equivalent to the assertion that for each $x \in X$, the set $\{\tau_1^n x, -\infty < n < \infty\}$ is dense in X. By the foregoing lemma, the image of this set under π_n is dense in X_n for each n. By the definition of the topology in X, it follows that $\{\tau_1^n x\}$ is dense in X.

Define a metric in X by setting

$$d(x, x') = \max_n \{2^{-n} | \zeta_n - \zeta'_n | \tag{13.4}$$

where $x = (\zeta_1, \zeta_2, \cdots, \zeta_n, \cdots)$ and $x' = (\zeta'_1, \zeta'_2, \cdots, \zeta'_n, \cdots)$. Note that if we consider X as a group with respect to pointwise multiplication, then this metric is invariant under group rotation. Using this metric, $F(x, x')$ is defined as in § 5.

LEMMA 13.4. *Let x, x' be two points in X with $\pi_n(x) = \pi_n(x')$ for some n. There exists a sequence of points $\{x_k\}$ with $\pi_{n-1}(x_k) = \pi_{n-1}(x) = \pi_{n-1}(x')$ and $F(x_k, x) \to 0$, $F(x_k, x') \to 0$, as $k \to \infty$.*

Proof. Let

$$x = (\zeta_1^0, \zeta_2^0, \cdots, \zeta_{n-1}^0, \zeta_n, \zeta_{n+1}, \cdots),$$
$$x' = (\zeta_1^0, \zeta_2^0, \cdots, \zeta_{n-1}^0, \zeta_n, \zeta'_{n+1}, \zeta'_{n+2}, \cdots).$$

We shall choose x_k of the form $(\zeta_1^0, \zeta_2^0, \cdots, \zeta_{n-1}^0 \zeta_n e^{i\beta_k}, \zeta_{n+1}, \zeta_{n+2}, \cdots)$, where β_k is an irrational multiple of π and $|\beta_k| < 2^{-k}$. Then

$$d(x_k, x) = 2^{-n} | e^{i\beta_k} - 1 | < 2^{-n-k} \to 0 \text{ as } k \to \infty.$$

A fortiori, $F(x_k, x) \to 0$. We claim that we also have $F(x_k, x') \to 0$. More precisely, we shall show that $F(x_k, x') < 2^{-n-k}$.

Suppose that $u = (\xi_1, \cdots, \xi_n, \cdots)$ and $u' = (\xi'_1, \cdots, \xi'_n, \cdots)$ are two points of X. Considering X as a group, we may form the quotient $u/u' = (\xi_1/\xi'_1, \cdots, \xi_n/\xi'_n)$. We notice that the quotient of the transforms $\tau_1 u$ and $\tau_1 u'$ is a function of u/u':

$$\tau_1 u/\tau_1 u' = (\xi_1/\xi'_1, \xi_1/\xi'_1 \cdot \xi_2/\xi'_2, \cdots, \xi_{n-1}/\xi'_{n-1} \cdot \xi_n/\xi'_n, \cdots) = \sigma(u/u')$$

where

(13.5) $$\sigma(\xi_1, \xi_2, \cdots, \xi_n, \cdots) = (\xi_1, \xi_1\xi_2, \cdots, \xi_{n-1}\xi_n, \cdots).$$

It follows that

(13.6) $$F(u, u') = \inf d(\tau_1^n u, \tau_1^n u') = \inf d(\tau_1^n u/\tau_1^n u', 1) = \inf d(\sigma^n(u/u'), 1)$$

where 1 denotes the identity element of the group X. We shall apply (13.6) to evaluate $F(x_k, x')$

We have

$$w_k = x_k/x' = (1, 1, \cdots, 1, e^{i\beta_k}, \zeta_{n+1}/\zeta'_{n+1}, \zeta_{n+2}/\zeta'_{n+2}, \cdots)$$

where $e^{i\beta_k}$ occurs in the n-th position. Consider a general element of the form

(13.7) $$w = (1, 1, \cdots, 1, e^{i\beta_k}, \lambda_1, \lambda_2, \lambda_3, \cdots).$$

then

$$\sigma w = (1, 1, \cdots, 1, e^{i\beta_k}, e^{i\beta_k}\lambda_1, \lambda_1\lambda_2, \lambda_2\lambda_3, \cdots) = (1, 1, \cdots, 1, e^{i\beta_k}, \lambda'_1, \lambda'_1, \lambda'_3, \cdots)$$

where the transformation $(\lambda_1, \lambda_2, \lambda_3, \cdots) \to (\lambda'_1, \lambda'_2, \lambda'_3, \cdots)$ is that of (13.2) with $\alpha = \beta_k$. Now apply Lemma 13.3. We find that the iterates $\{\sigma^m w_k, -\infty < m < \infty\}$ are dense in the set of elements of the form (13.7). In particular these come arbitrarily close to the element

$$w^* = (1, 1, \cdots, 1, e^{i\beta_k}, 1, 1, \cdots)$$

which differs from 1 only in the n-th entry. Then

$$\inf d(\sigma^m w_k, 1) \leqq \inf d(\sigma^m w_k, w^*) + d(w^*, 1) = d(w^*, 1) \\ = 2^{-n} \mid e^{i\beta_k} - 1 \mid < 2^{-n-k}.$$

This proves that $F(x_k, x') < 2^{-n-k}$, and our lemma is established.

THEOREM 13.3. *(X_n, T) is the maximal isometric extension of (X_{n-1}, T) in (X, T) and the order of (X, T) is ω.*

Proof. Let $(Y, T) = \pi'(X, T)$ be the maximal isometric extension of (X_{n-1}, T) in (X, T). Then $(X_n, T) \prec (Y, T)$. If these two subflows are not identical, then there exist two points x, x' with $\pi_n(x) = \pi_n(x')$ but $\pi'(x) \neq \pi'(x')$. Form the sequence $\{x_k\}$ in accordance with the preceding lemma. That is $\pi_{n-1}(x_k) = \pi_{n-1}(x) = \pi_{n-1}(x')$ and $F(x, x_k) \to 0$ $F(x', x_k) \to 0$. Suppose $\rho(y_1, y_2)$ is the metric for the fibres of Y over X_{n-1}, and set $R(x_1, x_2) = \rho(\pi' x_1, \pi' x_2)$. Since the points x_k and x are in the same fibre of X with respect to X_{n-1}, $R(x, x_k)$ is defined. Since R is invariant under T, $R(x, x_k) = R(\tau_1{}^m x, \tau_1{}^m x_k)$. Now $F(x, x_k) \to 0$ implies that for a sequence $\{m_k\}$, $d(\tau_1{}^{m_k} x, \tau_1{}^{m_k} x_k) \to 0$ and so also (since R is continuous) $R(\tau_1{}^{m_k} x, \tau_1{}^{m_k} x_k) \to 0$. Hence $R(x, x_k) \to 0$. Similarly, $R(x', x_k) \to 0$. But then $R(x, x') \leqq R(x, x_k) + R(x_k, x') \to 0$ and $R(x, x') = 0$. This implies that $\pi' x = \pi' x'$ contrary to our supposition. It follows that (X_n, T) is the maximal isometric extension of (X_{n-1}, T) in (X, T). The second assertion of the theorem follows from the first according to Theorem 13.2. This completes the proof of this theorem.

UNIVERSITY OF MINNESOTA.

REFERENCES.

[1] L. Auslander, L. Green, and F. Hahn, *Flows on Homogeneous Spaces*, Annals of Mathematics Study No. 53, Princeton.

[2] L. Auslander, F. Hahn, and L. Markus, "Topological dynamics on nilmanifolds," *Bulletin of the American Mathematical Society*, vol. 67 (1961), 298-299.

[3] N. Dunford and J. Schwartz, *Linear Operators*, Interscience Publishers, New York (1958).

[4] R. Ellis, "Distal transformation groups," *Pacific Journal of Mathematics*, vol. 8 (1958), pp. 401-405.

[5] ———, "Locally compact transformation groups," *Duke Mathematical Journal*, vol. 24 (1957), pp. 119-125.

[6] H. Furstenberg, "Strict ergodicity and transformations of the torus," *American Journal of Mathematics*, vol. 83 (1961), pp. 573-601.

[7] W. H. Gottschalk and G. A. Hedlund, *Topological Dynamics*, American Mathematical Society Colloquium Publications, vol. 83 (1961), pp. 573-601.

[8] F. Riesz and B. Sz-Nagy, *Leçons d'Analyse Fonctionnelle*, Budapest (1952).

SPACES WITH INVOLUTION AND BUNDLES OVER P^n.*[1]

By J. LEVINE.

In this paper we will present a general technique for studying fibre bundles over the real projective spaces P^n. Our results will, in particular, enable us to give a complete description of the vector bundles over P^n for $n \leqq 4$.

We begin, in part I, with a general classification theorem for bundles over P^n. The analogy to the Feldbau Classification of bundles over spheres will be clear. But, as in the latter, this does not remove the difficulty of computations in particular cases. We are led to another problem.

For any space with a base point and on involution which leaves the base point fixed, we will define *equivariant homotopy sets*—which will usually be groups. A calculation of these sets, in the case when they are *not* groups, is our new problem. In part II we develop some machinery for this purpose.

In part III these techniques are applied to the study of vector bundles. We obtain the results mentioned above; these are summarized in 7.7.

It might be worth mentioning that the theory of part II is developed primarily as a tool in the problem of classifying fibre bundles. We have not here considered its possible independent interest in the study of spaces with involutions. This will be the subject of a future paper.

Part I.

1. Preliminaries. 1.1. Throughout Part I, G will denote a *compact Lie group* with identity element e. We will denote a *principal G-bundle* by:

$$\mathcal{B} = (B, p, X, G)$$

where B is the total space, X the base space, p the projection and G the structural group. We refer the reader to [6] for the necessary definitions. If

$$\mathcal{B}' = (B', p', X', G)$$

is another principal G-bundle, a *bundle map* $\mathcal{B} \to \mathcal{B}'$ will be denoted by a pair of mappings (ϕ, φ) satisfying the following commutative diagram:

* Received August 1, 1962.

[1] This paper is essentially the major part of a doctoral dissertation presented to Princeton University, 1962. I am indebted to my adviser, N. E. Steenrod, for his assistance and encouragement.

$$\begin{tikzcd} B \arrow[r, "\phi"] & B' \\ X \arrow[u, "p"] \arrow[r, "\varphi"] & X' . \arrow[u, "p'"'] \end{tikzcd}$$

Also see [6] for the definition.

1.2. We will denote n-dimensional real projective space by P^n, the unit n-sphere by S^n and the closed unit n-disk by D^n. Let $i\colon D^n \to S^n$ be a homeomorphism of D^n onto the upper hemisphere so that $i \mid S^{n-1}$ is the inclusion. Let $h\colon D^n \to P^n$ be the composition of i with the standard double covering $S^n \to P^n$. Let $\mu\colon S^n \to S^n$ denote the antipodal map.

1.3. A map $f\colon S^{n-1} \to G$ called *equivariant* or an *e-map* if:

$$f\mu(x) = f(x)^{-1} \text{ for } x \in S^{n-1}.$$

A homotopy $f_t\colon S^{n-1} \to G$ is called an *equivariant homotopy* (*e-homotopy*) if each f_t is an e-map. An e-homotopy $\{f_t\}$, $0 \leqq t \leqq 1$, is a *ce-homotopy* from f to g if $f = af_0a^{-1}$ and $g = bf_1b^{-1}$ for some $a, b \in G$.

2. The classification theorem. 2.1. Let $\mathfrak{B} = (B, p, P^n, G)$ be a G-bundle. Since D^n is contractible, there is a map H satisfying the following commutative diagram:

$$\begin{tikzcd} & B \arrow[d, "p"] \\ D^n \arrow[ur, "H"] \arrow[r, "h"] & P^n . \end{tikzcd}$$

We define a map $f\colon S^{n-1} \to G$ by the following equation, using the right action of G on B:

$$H(x) = H\mu(x)f(x) \qquad \text{for } x \in S^{n-1}.$$

It is straightforward to check that f is *equivariant*; we call f the *characteristic map of* $\mathfrak{B}$ associated with H.

2.2. Proposition. *Any two characteristic maps of* $\mathfrak{B}$ *are ce-homotopic.*

Proof. Let $f, f'\colon S^{n-1} \to G$ be characteristic maps of $\mathfrak{B}$ associated with $H, H'\colon D^n \to B$. We define $g\colon D^n \to G$ by:

$$H'(x) = H(x)g(x).$$

We can immediately verify:

$$f'(x) = g\mu(x)^{-1}f(x)g(x) \qquad \text{for } x \in S^{n-1}.$$

Now define $f_t\colon S^{n-1} \to G$, for $0 \leqq t \leqq 1$, by:

$$f_t(x) = g(t\mu(x))^{-1}f(x)g(tx).$$

Clearly $\{f_t\}$ is a *ce*-homotopy from f to f'.

2.3. From 2.2, it follows that consideration of the characteristic map associates to every *G-bundle* $\mathcal{B}$ a *ce*-homotopy class of maps $S^{n-1} \to G$, which we shall denote by $\chi(\mathcal{B})$. We now state the classification theorem.

THEOREM. *The correspondence* χ *is bijective.*

The proof of this theorem will occupy the next two sections.

2.4. We first establish *surjectivity*. Let $f\colon S^{n-1} \to G$ be equivariant; we define a G-bundle $\mathcal{B}_f = (B_f, p_f, P^n, G)$ as follows.

Consider the trivial G-bundle $\mathcal{T} = (D^n \times G, p, D^n, G)$ and sub-bundle $\mathcal{T}' = (S^{n-1} \times G, p', S^{n-1}, G)$. Define a bundle map $(M, \mu)\colon \mathcal{T}' \to \mathcal{T}'$ by:

$$M(x, y) = (\mu(x), f(x)y) \qquad \text{for } x \in S^{n-1}, y \in G.$$

Recall that μ is the antipodal map. Since f is equivariant and μ is an involution, M is an involution. We now construct $\mathcal{B}_f$ by collapsing $\mathcal{T}$ under (M, μ). More precisely, define B_f to be $D^n \times G$ collapsed under M and p_f the projection induced from p, *i. e.*, so that the following diagram is commutative:

$$\begin{tikzcd} D^n \times G \arrow[r, "H'"] \arrow[d, "p"] & B_f \arrow[d, "p_f"] \\ D^n \arrow[r, "h"] & P^n \end{tikzcd}$$

where H' is the collapsing map. A right action of G on B_f is defined by demanding that H' be a G-map. It is clear that G acts without fixed point and p_f collapses orbits. We can apply a theorem of Gleason [2] and conclude that $\mathcal{B}_f$ is a G-bundle.

Define $H\colon D^n \to B_f$ by $H(x) = H'(x, e)$; clearly $p_f H = h$. It is a straightforward exercise to check that f is, indeed, the characteristic map of $\mathcal{B}_f$ associated with H. Surjectivity follows immediately.

2.5. We now verify *injectivity*. For $i = 0, 1$, let $\mathcal{B}_i = (B_i, p_i, P^n, G)$ be a G-bundle with characteristic map f_i associated with H_i. Suppose f_0 and

f_1 are *ce*-homotopic; we will show B_0 and B_1 are equivalent. It clearly suffices to deal with two cases.

(i) $f_0 = a^{-1}f_1 a$ *for some* $a \in G$.

(ii) f_0 *and* f_1 *are e-homotopic.*

Case (i). We define an equivalence $(E,1): \mathcal{B}_0 \to \mathcal{B}_1$, when 1 is an identity map, by the equation:

$$E(H_0(x)y) = H_1(x)ay \qquad \text{for } x \in D^n, y \in G.$$

If $H_0(x)y = H_0(x')y'$ and $x \neq x'$, then $x \in S^{n-1}$. A straightforward calculation shows that:

$$E(H_0(x)y) = E(H_0(x')y').$$

Therefore E is well-defined. It is easily checked that $(E,1)$ is a bundle map; this proves (i).

Case (ii). By (i), we may assume $\mathcal{B}_i = \mathcal{B}_{f_i}$, as defined in 2.4. Let $f_t: S^{n-1} \to G$, $0 \leqq t \leqq 1$, be an e-homotopy from f_0 to f_1. We will construct a bundle $\mathcal{B} = (B, p, I \times P^n, G)$ such that $\mathcal{B} \mid (i \times P^n) = \mathcal{B}_i$ for $i = 0, 1$. The construction is analogous to that in 2.4.

Consider the trivial bundle $\mathcal{T} = (I \times D^n \times G, p', I \times D^n, G)$ and subbundle $\mathcal{T}' = (I \times S^{n-1} \times G, p'', I \times S^{n-1}, G)$. Define a bundle involution $(M', \mu'): \mathcal{T}' \to \mathcal{T}'$ by:

$$M'(t,x,y) = (t, \mu(x), f_t(x)y)$$

$$\mu'(t,x) = (t, \mu(x)) \qquad \text{for } t \in I, x \in S^{n-1}, y \in G.$$

We obtain $\mathcal{B}$ by collapsing $\mathcal{T}$ under (M', μ'). Proceeding as in 2.4, we verify that $\mathcal{B}$ is a G-bundle. It is obvious that $\mathcal{B} \mid (i \times P^n) = \mathcal{B}_i$ for $i = 0, 1$.

This completes the proof of the classification theorem.

2.6. Let $\pi_{n-1}{}^0(G)$ denote the e-homotopy classes of maps $S^{n-1} \to G$. We can define a group of $\pi_0(G)$ on $\pi_{n-1}{}^0(G)$ by conjugation: If $a \in G$ and $f: S^{n-1} \to G$ is an e-map, consider the e-map afa^{-1}. If $\{a_t\}$ is a path in G, then $\{a_t f a_t^{-1}\}$ is an e-homotopy; thus the action is well-defined.

In this framework, *the ce-homotopy classes of e-maps $S^{n-1} \to G$ can be identified, in a natural manner, with the orbits of $\pi_{n-1}{}^0(G)$ under the action* of $\pi_0(G)$. This gives a more useful interpretation of the classification of G-bundles over P^n.

Theorem 2.3 allows us to define an operation Θ: if $\beta \in \pi_{n-1}{}^0(G)$, let $\Theta(\beta)$ be the G-bundle such that $\chi\Theta(\beta)$ is the orbit of β.

We will make use of Θ in part III.

Part II.

3. Homotopy in spaces with an Involution. 3.1. In this section we initiate a study of the $\pi_{n-1}{}^0(G)$, introduced in 2.6. Since the only structure on G in which we are interested is inversion, it is useful to expand our considerations to any space X with involution T. Thus we define $\pi_n{}^0(X;T)$.

If the space X has a base point x_0 fixed under T, we may define still more homotopy invariants of (X, x_0, T). Let T_i be a reflection of S^n about some S^{i-1}; the equivariant homotopy classes $(S^n, S^{i-1}) \to (X, x_0)$ form a group if $i > 0$. If $i = 0$ this gives the $\pi_n{}^0(X;T)$ again; these are not groups, but there is a group action by $\pi_n(X)$.

3.2. Let S^n and D^{n+1} denote the unit sphere and closed unit disk in Euclidean $(n+1)$-space E^{n+1}. The points of S^n will often be represented by pairs (x, y) or triples (x, t, y) where $x \in E^k$, $t \in E^1$ and $y \in E^r$.

Let T_i, $0 \leqq i \leqq n$, denote the *reflection of S^n about the first i coordinates*:

$$T_i(x, y) = (x, -y) \qquad \text{if } x \in E^i.$$

We agree to identify S^{i-1} with the unit sphere on the first i coordinates in E^{n+1}. Clearly, *S^{i-1} is the fixed point set of T_i*. Note that T_0 is the antipodal map.

Let X be a *connected* space with an involution T. We assume X is *n-simple for all n*, i.e. the fundamental group acts trivially on all homotopy groups of X. Thus $\pi_n(X)$ is well-defined. Let x_0 be a basepoint of X such that $T(x_0) = x_0$; we say $(X, x_0; T)$ is a *space with involution*.

3.3. Let $F_n{}^i(X, x_0; T)$ denote the space of mappings $f\colon (S^n, S^{i-1}) \to (X, x_0)$ satisfying $fT_i = Tf$, with the compact-open topology. The constant map into x_0, which we also denote by x_0, is a basepoint of $F_n{}^i(X, x_0; T)$. If $i = 0$, the base-point x_0 is unnecessary and we may, more generally, define $F_n{}^0(X;T)$.

The set of arc-components of $F_n{}^i(X, x_0; T)$ will be denoted by $\pi_n{}^i(X, x_0; T)$. The arc-component of x_0, denoted also by x_0, is a base-point. If $i = 0$, we may define $\pi_n{}^0(X;T)$. Note that, if $X = G$ and T is inversion, $\pi_n{}^0(X;T) = \pi_n{}^0(G)$, as in 2.6.

In any discussion where $(X, x_0; T)$ is fixed, we will use the abbreviations F_n^i and π_n^i.

3.4. It will be convenient to have notation for the hemispheres and equators along different coordinate axes in S^n. Let $D_{i,+}^n$, $D_{i,-}^n$ and S_i^{n-1} respectively, denote the set of points (x, t, y) for $x \in E^i$, where $t \geqq 0$, $t \leqq 0$ and $t = 0$, respectively. We identify S_i^{n-1} with S^{n-1} by means of the map $e_i \colon S^{n-1} \to S^n$ defined by:

$$e_i(x, y) = (x, 0, y) \qquad \text{if } x \in E^i.$$

The following lemma will be useful.

LEMMA. *If $0 \leqq i \leqq n-1$ and $f \colon D_{i,+}^n \to X$ satisfies $fe_i \in F_{n-1}^i$, then f extends uniquely to a map in F_n^i.*

Proof. We extend f over $D_{i,-}^n$ by:

$$f \mid D_{i,-}^n = TfT_i \mid D_{i,-}^n.$$

This is well-defined because $T_i(D_{i,-}^n) = D_{i,+}^n$ and is consistent on S_i^{n-1} because $fe_i \in F_{n-1}^i$. Clearly $fT_i = Tf$ and $f(S^{i-1}) = x_0$, since $S^{i-1} \subset S_i^{n-1}$. Uniqueness is obvious.

3.5. It is clear that the involution T_{i+1} on S^{n+1} may be regarded as the "suspension" of the involution T_i on S^n. This fact may be expressed by a *natural identification of F_{n+1}^{i+1} with the space of loops of F_n^i based at x_0.*

Let $S_i^n(t)$, $-1 \leqq t \leqq 1$, denote the family of parallels along the $(i+1)$-st axis in S^{n+1} i.e. $S_i^n(t)$ is the set of points (x, t, y), where $x \in E^i$. Unless $t = \pm 1$, $S_i^n(t)$ is an n-sphere; if $t = \pm 1$, it is a point. In any case we represent it as an n-sphere by the mapping $u_t^i \colon S^n \to S^{n+1}$ defined by:

$$u_t^i(x, y) = (c_t x, t, c_t y) \qquad \text{if } x \in E^i,$$

where $c_t = (1 - t^2)^{\frac{1}{2}}$. If $f \in F_{n+1}^{i+1}$, it is easy to check that $f_t = fu_t^i \in F_n^i$ for $-1 \leqq t \leqq 1$. Clearly $f_1 = f_{-1} = x_0$ since $S_i^n(\pm 1) \in S^i$. Thus $\{f_t\}$ is a loop in F_n^i based at x_0. Conversely, if $\{f_t\}$ is a loop in F_n^i based at x_0, we can define $f \in F_{n+1}^{i+1}$ by $fu_t^i = f_t$ for $-1 \leqq t \leqq 1$.

Thus every F_n^i is a loop space if $i > 0$ and, consequently, has a natural H-space structure. Suppose $f_1, f_2 \in F_n^i$; define $f \colon S^n \to X$ as follows:

$$f(x, t, y) = \begin{cases} f_1(c'_t x,\ 2t - 1, c'_t y) & \text{if } t \geqq 0, \\ f_2(c''_t x, 2t + 1, c''_t y) & \text{if } t \leqq 0, \end{cases}$$

where $x \in E^{i+1}$, $c'_t = 2(t/(1+t))^{\frac{1}{2}}$, $c''_t = 2(t/(t-1))^{\frac{1}{2}}$. This is well-defined when $t = 0$ because $f_1(0, -1, 0) = f_2(0, 1, 0) = x_0$. It is straightforward to

check that $f \in F_n{}^i$. In the H-space multiplication inherited from loop multiplication, f is the product of f_1 and f_2.

Finally, we can conclude that, for $i > 0$, the $\pi_n{}^i$ inherit a *group structure.* If $i = 0$, there is no analogous structure.

3.6. *Remark.* Let A be a topological space with base point a_0. Let ΩA be the loop space of A based at a_0 and T the involution of ΩA defined by reversing the orientation of loops. Let s^iP^k denote the i-fold reduced suspension of real projective k-space. Then there is a natural homeomorphism between $\dot{F}_n{}^i(\Omega A, a_0; t)$ and the space of mappings $(s^iP^{n-i+1}, x) \to (A, a_0)$, where x is a base point of P^{n-i+1}. If $i > 0$, "track multiplication" is defined on this space and corresponds to the multiplication in $F_n{}^i(\Omega A, a_0; T)$.

If $A = B_G$ the classifying space of the group G, this means that G-bundles over P^n are classified by $\pi_{n+1}{}^0(\Omega B_G; T)$ after an appropriate collapsing under action of $\pi_0(\Omega B_G) \approx \pi_1(B_G)$. This is closely related to the results of Part I.

3.7. There is another algebraic structure on the $\pi_n{}^i(X, x_0, T)$ for all $i \geqq 0$. We will show that $\pi_n = \pi_n(X)$ *is a group of operators on the* $\pi_n{}^i$. First we derive a standard lemma.

Lemma. *Suppose X is an n-simple space and $f_0, f_1\colon S^n \to X$ are homotopic. Furthermore, let $g_t\colon D_{i,-}{}^n \to X$, $0 \leqq t \leqq 1$, be a homotopy of $g_0 = f_0 \mid D_{i,-}{}^n$ to $g_1 = f_1 \mid D_{i,-}{}^n$. Then $\{g_t\}$ extends to a homotopy of f_0 to f_1.*

Proof. It suffices to prove the following statement.

(*) Suppose $M \subset L \subset K$ are subcomplexes and complexes and M is a deformation retract of L. Let $f\colon L \to X$, for any space X. If $f \mid M$ is extendible over K, so is f.

To see that the lemma follows from (*), let $y \in D_{i,-}{}^n$ and set $K = I \times S^n$, $L = (I \times D_{i,-}{}^n) \cup (\dot{I} \times S^n)$ and $M = (I \times y) \cup (\dot{I} \times S^n)$. Define f by:

$$f(t,x) = \begin{cases} f_t(x) & \text{if } t \in \dot{I}, x \in S^n. \\ g_t(x) & \text{if } t \in I, x \in D_{i,-}{}^n. \end{cases}$$

Now $f \mid M$ is extendible over K because f_0 is homotopic to f_1 and X is n-simple. That f is extendible over K is the conclusion of the lemma.

Proof of (*). Let $r_t\colon L \to L$ be a homotopy such that $r_0 =$ identity, $r_1(L) = M$ and $r_t \mid M$ is the inclusion, for $0 \leqq t \leqq 1$. Since $f \mid M$ is extendible over K, so is fr_1. But then, by the homotopy extension theorem, $f = fr_0$ is extendible over K. This proves (*) and, therefore, the lemma.

3.8. Let $\alpha \in \pi_n$, $\beta \in \pi_n{}^i$ and define $r_i\colon S^n \to S^n$ to be the reflection along the $(i+1)$-st axis, i.e.

$$r_i(x,t,y) = (x,-t,y) \qquad \text{if } x \in E^i.$$

Choose representatives $f_0\colon S^n \to X$ and $g_0 \in F_n{}^i$ of α and β respectively, so that:

$$g_0 r_i \mid D_{i,-}{}^n = f_0 \mid D_{i,-}{}^n.$$

By 3.4, we may extend $f_0 \mid D_{i,+}{}^n$ to $h_0 \in F_n{}^i$, because $f_0 e_i = g_0 e_i \in F_{n-1}{}^i$. Let $\gamma \in \pi_n{}^i$ be the homotopy class of h_0; we shall show that γ depends only on α and β.

Let $f_1\colon S^n \to X$ and $g_1 \in F_n{}^i$ be any other representatives of α and β, respectively, such that:

$$g_1 r_i \mid D_{i,-}{}^n = f_1 \mid D_{i,-}{}^n.$$

Then we can define $h_1 \in F_n{}^i$, as above. Let $\{g_t\} \in F_n{}^i$ be a homotopy from g_0 to g_1. By 3.7 there is a homotopy $\{f_t\}$ from f_0 to f_1 such that:

$$g_t r_i \mid D_{i,-}{}^n = f_t \mid D_{i,-}{}^n \qquad \text{for } 0 \leqq t \leqq 1.$$

Then, from f_t and g_t, we can construct $h_t \in F_n{}^i$ for $0 \leqq t \leqq 1$. It is clear that $\{h_t\}$ is a homotopy in $F_n{}^i$ from h_0 to h_1; thus h_1 also represents γ.

We denote the element γ by $\alpha \cdot \beta$.

3.9. We now show that the operation $(\alpha,\beta) \to \alpha\cdot\beta$ defines a *group action of π_n on $\pi_n{}^i$*. Suppose α_1 and $\alpha_2 \in \pi_n$ and $\beta \in \pi_n{}^i$ are represented by $f_1, f_2\colon S^n \to X$ and $g \in F_n{}^i$, respectively, such that:

$$f_2 \mid D_{i,-}{}^n = g r_i \mid D_{i,-}{}^n$$
$$f_1 \mid D_{i,-}{}^n = f_2 r_i \mid D_{i,-}{}^n.$$

Then $\alpha_1 \cdot \alpha_2$ is represented by $f\colon S^n \to X$ defined by:

$$f \mid D_{i,+}{}^n = f_1 \mid D_{i,+}{}^n$$
$$f \mid D_{i,-}{}^n = f_2 \mid D_{i,-}{}^n = g r_i \mid D_{i,-}{}^n$$

and $\alpha_2 \cdot \beta$ is represented by $h \in F_n{}^i$ defined by:

$$h \mid D_{i,+}{}^n = f_2 \mid D_{i,+}{}^n$$

and 3.4. It then follows that:

$$h r_i \mid D_{i,-}{}^n = f_2 r_i \mid D_{i,-}{}^n = f_1 \mid D_{i,-}{}^n,$$

Now, it is straightforward to check that $\alpha_1 \cdot (\alpha_2 \cdot \beta)$ and $(\alpha_1 \cdot \alpha_2) \cdot \beta$ are both represented by $h' \in F_n{}^i$ defined by:

$$h' \mid D_{i,+}{}^n = f_1 \mid D_{i,+}{}^n$$

and 3.4. This proves that $(\alpha\cdot\beta)\to\alpha\cdot\beta$ is a group action.

4. An exact sequence.

4.1. Let $(X, x_0; T)$ be a space with involution. We define:

$$\phi_i\colon \pi_n\to\pi_n{}^i$$
$$\psi_i\colon \pi_n{}^i\to\pi_{n-1}{}^i$$
$$\rho_i\colon \pi_n{}^i\to\pi_n.$$

The main result of this section (4.3) will be an exact sequence involving these functions.

Let $\beta\in\pi_n{}^i$ be represented by $g\in F_n{}^i$ and let $\alpha\in\pi_n$. Then $\phi_i(\alpha) = \alpha\cdot x_0$, $\psi_i(\beta)$ is represented by $ge_i\in F_{n-1}$ and $\rho_i(\beta)$ is represented by $g\colon S^n\to X$.

4.2. If $i>0$, ϕ_i, ψ_i and ρ_i are homomorphisms. That ρ_i is a homomorphism follows from the fact that the H-space structure of $F_n{}^i$, as defined in 3.5, agrees with the group structure of π_n. The function ψ_i *is* a homomorphism because multiplication is along the i-th coordinate axis while ψ_i is defined by restricting mappings to the equator along the $(i+1)$-st coordinate axis.

That ϕ_i is a homomorphism follows from the more general formula, for $i>0$:

$$(\alpha_1\cdot\alpha_2)\cdot(\beta_1\cdot\beta_2)=(\alpha_1\cdot\beta_1)\cdot(\alpha_2\cdot\beta_2) \qquad \text{if } \alpha_j\in\pi_n, \beta_j\in\pi_n{}^i.$$

This formula follows from the fact that multiplication is defined along the i-th coordinate axis while the group action is defined along the $(i+1)$-st coordinate axis.

From the formula it follows that, if $i>0$, $\alpha\cdot\beta=\phi_i(\alpha)\cdot\beta$. Thus, the group action of π_n on $\pi_n{}^i$ is determined by the group structure on $\pi_n{}^i$ and ϕ_i, if $i>0$. Only if $i=0$ is the group action significant.

4.3. THEOREM. *The following sequence is exact:*

$$\cdots\to\pi_n{}^{i+1}\xrightarrow{\rho_{i+1}}\pi_n\xrightarrow{\phi_i}\pi_n{}^i\xrightarrow{\psi_i}\pi_{n-1}{}^i\xrightarrow{\rho_i}\pi_{n-1}\to\cdots$$

$$\cdots\to\pi_{n-i}{}^1\xrightarrow{\rho_1}\pi_{n-i}\xrightarrow{\phi_0}\pi_{n-i}{}^0\xrightarrow{\psi_0}\pi_{n-i-1}{}^0\xrightarrow{\rho_0}\pi_{n-i-1}$$

where exactness at the right end of the sequence is defined by:

(a) *Image* $\rho_1=\phi_0{}^{-1}(x_0)$.

(b) $\psi_0{}^{-1}(\beta)$ *is an orbit of the action of* π_{n-i} *on* $\pi_{n-i}{}^0$ *for* $\beta\in\pi_{n-i-1}{}^0$.

(c) *Image* $\psi_0=\ker\rho_0$.

Remarks. (1) If T has *no fixed point*, $\pi_n{}^0$ is still defined together with the group action of π_n on $\pi_n{}^0$ and the functions ψ_0 and ρ_0. In this case, (b) and (c) are still valid.

(2) Suppose, as in 3.6, that $X = \Omega A$, where A is a space with base point a_0, and x_0 is the constant loop at a_0. Let $f\colon S^k \to P^k$ be the double covering; in [5], Puppe associates with the map f and space A, an exact sequence:

$$\begin{aligned} \cdots \to \pi(s^{i+1}P^k, A) \to \pi(s^{i+1}S^k, A) \to \pi(s^iC_f, A) \\ \to \pi(s^iP^k, A) \to \pi(S^iS^k, A) \to \cdots \end{aligned}$$

$$\cdots \to \pi(sP^k, A) \to \pi(sS^k, A) \to \pi(C_f, A) \to \pi(P^k, A) \to \pi(S^k, A)$$

where $\pi(B_1, B_2)$ is the set of homotopy classes of base-preserving maps $B_1 \to B_2$ and C_f is the "mapping cone" of f. It is easy to check that C_f is homeomorphic to P^{k+1}. Using the remark of 3.6, one can show that the exact sequence of Puppe coincides with that of the theorem.

4.4. *Proof of Theorem.* It clearly suffices to prove the following statements for all $i \geqq 0$:

(a) *Image* $\rho_{i+1} = \phi_i^{-1}(x_0)$.
(b) $\psi_i^{-1}(\beta)$ *is an orbit of the action of* π_n *on* $\pi_n{}^i$ *for* $\beta \in \pi_{n-1}{}^i$.
(c) *Image* $\psi_i = \ker \rho_i$.

The proof of these statements will occupy the next three sections.

4.5. We first prove (a) of 4.4. Let $g' \in F_n{}^{i+1}$ represent $\beta \in \pi_n{}^{i+1}$; we will show $\phi_i\rho_{i+1}(\beta) = x_0$. Let $\{g'_t\} \in F_{n-1}{}^i$ be the loop defined by $g'_t = gu_t{}^i$, as in 3.5, for $-1 \leqq t \leqq 1$. Reparametrize this loop by any map $w\colon (D^1, S^0) \to (D^1, S^0)$ such that:

$$w(t) = \begin{cases} -1 & \text{if } t \leqq 0, \\ 1 & \text{if } t = 1. \end{cases}$$

This defines a loop $\{g_t\}$, where $g_t = g'_{w(t)}$, and corresponding $g \in F_n{}^{i+1}$, defined by $gu_t{}^i = g_t$. Any homotopy of w to the identity, modulo S^0, induces a homotopy between g and g' in $F_n{}^{i+1}$. Therefore g represents β and, clearly, $g(D_{i,-}{}^n) = x_0$.

Now $\phi_i\rho_{i+1}(\beta)$ is represented by the extension of $g \mid D_{i,+}{}^n$ to an element of $F_n{}^i$. To show $\phi_i\rho_{i+1}(\beta) = x_0$, by 3.4 it suffices to construct a homotopy $h_s\colon D_{i,+}{}^n \to X$, $0 \leqq s \leqq 1$, satisfying:

$$\begin{aligned} &h_0 = g \mid D_{i,+}{}^n \\ &h_1 = x_0 \\ &h_se_i \in F_{n-1}{}^i \qquad\qquad \text{if } 0 \leqq s \leqq 1. \end{aligned}$$

Note that $g \mid D_{i,+}^n$ corresponds to the *path* in F_{n-1}^i given by $\{g_t\}$ for $0 \leqq t \leqq 1$. For each s between 0 and 1, we consider the path in F_{n-1}^i defined by cutting $\{g_t\}$ down to the interval $s \leqq t \leqq 1$. More precisely define $h_{s,t} \in F_{n-1}^i$ by:

$$h_{s,t} = g_{s+(1-s)t}.$$

Then we define h_s by:

$$h_s u_t^i = h_{s,t}.$$

It is easy to check that $\{h_s\}$ is the desired homotopy since $h_s e_i = h_{s,0} = g_s \in F_{n-1}^i$.

4.6. To complete the proof of (a), we must show that, if $\alpha \in \pi_n^i$ and $\phi_i(\alpha) = x_0$, then $\alpha \in \operatorname{Image} \rho_{i+1}$. Let $f\colon (S^n, D_{i,-}^n) \to (X, x_0)$ represent α. That $\phi_i(\alpha) = x_0$ means there is a homotopy $f_t\colon D_{i,+}^n \to X$, $0 \leqq t \leqq 1$, such that:

$$f_0 = f \mid D_{i,+}^n$$
$$f_1 = x_0$$
$$f_t e_i \in F_{n-1}^i.$$

Now define $g \in F_n^{i+1}$ by:

$$g u_t^i = \begin{cases} f_t e_i & \text{if } 0 \leqq t \leqq 1, \\ x_0 & \text{if } -1 \leqq t \leqq 0. \end{cases}$$

It is not hard to use the homotopy $\{f_t\}$ to construct a homotopy between f and g, stationary on $D_{i,-}^n$. Therefore $\alpha \in \operatorname{Image} \rho_{i+1}$.

4.7. We prove (b) of 4.4. It is clear that $\psi_i(\alpha \cdot \beta) = \psi_i(\beta)$, for any $\alpha \in \pi_n$, $\beta \in \pi_n^i$, since, in the construction of 3.8, the representative map of $\alpha \cdot \beta$ coincides with that of β on S_i^{n-1}.

Suppose $\beta_1, \beta_2 \in \pi_n^i$ and $\psi_i(\beta_1) = \psi_i(\beta_2)$. Let $g_j \in F_n^i$ represent β_j; then $g_1 e_i$ is homotopic to $g_2 e_i$ in F_{n-1}^i. By the homotopy extension theorem and 3.4, g_1 is homotopic in F_n^i to g'_1 satisfying $g'_1 e_i = g_2 e_i$. Define $f\colon S^n \to X$ by:

$$f \mid D_{i,+}^n = g_2 \mid D_{i,+}^n$$
$$f \mid D_{i,-}^n = g'_1 r_i \mid D_{i,-}^n.$$

We may check, using 3.8, that $\alpha \cdot \beta_1 = \beta_2$, if f represents α. This completes the proof of (b).

To prove (c), we merely note that $\beta \in \operatorname{Ker} \rho_i$ if and only if a representative of β is extendible over D^n. By 3.4, this is equivalent to $\beta \in \operatorname{Image} \psi_i$.

This completes the proof of 4.4, and, therefore, 4.3.

4.8. We conclude this section with two useful results incorporated in the following proposition. T_* will denote the involution of π_n induced by T.

PROPOSITION. (a) *If* $\alpha \in \pi_n$, $\beta \in \pi_n{}^i$, *then:*

$$\rho_i(\alpha \cdot \beta) = \rho_i(\beta) + \alpha + (-1)^{n-i+1} T_*(\alpha).$$

(b) *The homomorphism* $\phi_n \colon \pi_n \to \pi_n{}^n$ *is an isomorphism.*

Proof. To prove (a), let $f \colon S^n \to X$ represent α and $g \in F_n{}^i$ represent β, and we assume:

(1) $$f \mid D_{i,-}{}^n = g r_i \mid D_{i,-}{}^n.$$

Then, by 3.8, $\alpha \cdot \beta$ is represented by $h \in F_n{}^i$, where:

(2) $$h \mid D_{i,+}{}^n = f \mid D_{i,+}{}^n.$$

From (1) and (2) it is easy to derive the following:

$$\begin{aligned} h \mid D_{i,+}{}^n &= f \mid D_{i,+}{}^n \\ f \mid D_{i,-}{}^n &= g r_i \mid D_{i,-}{}^n \\ g \mid D_{i,-}{}^n &= (T f T_i) r_i \mid D_{i,-}{}^n \\ T f T_1 \mid D_{i,-}{}^n &= h \mid D_{i,-}{}^n. \end{aligned}$$

It follows from these that, in π_n, the element represented by h is the sum of the elements represented by f, g and TfT_i. Since $\deg T_i = (-1)^{n-i-1}$, this proves (a).

To prove (b) we notice that, by 3.4, $\pi_n{}^n$ is the set of homotopy classes of maps $(D_{i,+}{}^n, S_i{}^{n-1}) \to (X, x_0)$ and ϕ_n is the natural isomorphism with π_n.

5. Stability and double coverings.

5.1. We will first derive a theorem (5.2) relating the $\pi_n{}^i$ of a space with involution and an invariant subspace under connectivity assumptions on the pair. We then study the $\pi_n{}^i$ of a double covering of a space with involution (5.6).

For $j = 1, 2$ let $(X_j, x_j; T_j)$ be a space with involution (see 3.2). Suppose $f \colon (X_1, x_1) \to (X_2, x_2)$ satisfies $f T_1 = T_2 f$. Then we define:

$$f_\# \colon \pi_n{}^i(X_1, x_1; T_1) \to \pi_n{}^i(X_2, x_2; T_2)$$

in the natural manner. It is straightforward to check that $f_\#$ preserves the group action of 3.8 in the sense of:

$$(f_* \alpha) \cdot (f_\# \beta) = f_\#(\alpha \cdot \beta) \qquad \text{if } \alpha \in \pi_n(X_1), \beta \in \pi_n{}^i(X_1),$$

and $f_\#$ commutes with the functions of 4.1.

5.2. We will prove the following "stability" theorem.

THEOREM. *Suppose $(X, x_0; T)$ is a space with involution and A is a subspace of X containing x_0 and invariant under T such that $(A, x_0; T)$ is a space with involution. Let $i: A \to X$ be the inclusion and suppose (X, A) is k-connected. Then the induced function:*

$$i_{\#}: \pi_n{}^i(A, x_0; T) \to \pi_n{}^i(X, x_0; T)$$

(a) *is onto if $k \geqq n$.*

(b) *is one-one if $k \geqq n+1$.*

Proof. It suffices to prove the following statement:

(*) Let K be an n-complex with subcomplex L and X a space with subspace A. Suppose the finite group π acts on K (simplicially) and on X so that L and A are invariant. Suppose, in addition, that (X, A) is n-connected and π acts freely on K-L. Then any π-equivariant map $f: (K, L) \to (X, A)$ can be π-equivariantly deformed, modulo L, into A.

5.3. To see that the theorem follows from (*) we proceed as follows. Recall that T_i acts freely on $S^n - S^{i-1}$. To prove (a), let $K = S^n$, $L = S^{i-1}$ and $\pi = Z_2$ generated by T_i.

To prove (b), we apply (*) to a more complicated situation. Let $K = I \times S^n$, $L = (I \times S^{i-1}) \cup (\dot{I} \times S^n)$ and $\pi = Z_2$ generated by the involution $(t, x) \to (t, T_i(x))$; then (*) shows that any homotopy in $F_n{}^i(X, x_0; T)$ between two maps in $F_n{}^i(A, x_0; T)$ can be compressed into $F_n{}^i(A, x_0; T)$.

The proof of (*) is quite standard. We construct the deformation skeleton by skeleton, using the n-connectedness of (X, A) to solve the extension problems and the free action on K-L to maintain π-equivariance. We omit further details.

5.4. Let $(X, x_0; T)$ be a space with involution and $p: (Y, y_0) \to (X, x_0)$ a double-covering with covering transformation $t: Y \to Y$. Suppose we assume the condition:

$$(\dagger) \qquad T_* p_* \pi_1(Y) \subset p_* \pi_1(Y)$$

i.e. the subgroup of $\pi_1(X)$ corresponding to the covering p is invariant under T_*.

PROPOSITION. *There are precisely two involutions of Y-denoted by T' and T''—covering T $(pT' = pT'' = Tp)$. Moreover $T'T'' = T''T' = t$; in particular exactly one leaves y_0 fixed.*

Proof. This follows by elementary covering space arguments. By [3, p. 90], there are unique maps $f_j\colon Y \to Y$ for $j = 1, 2$, satisfying:

$$pf_j = Tp, \quad f_1(y_0) = y_0, \quad f_2(y_0) = t(y_0).$$

Now notice that $f_j^2 = f_jf_j$ satisfies:

$$pf_j^2 = p, \qquad f_j^2(y_0) = y_0;$$

therefore $f_j^2 =$ identity and f_j is an involution, by [3, p. 90]. Now, if $f = t$, f_1f_2 or f_2f_1, then:

$$pf = p, \qquad f(y_0) = t(y_0).$$

Therefore, by [3, p. 90], these are equal. The proposition follows by setting $f_1 = T'$, $f_2 = T''$.

Notation. Let T' be that involution of Y covering T satisfying $T'(y_0) = y_0$. Let $\bar{T}$ denote either T' or T''.
Note that $(Y, y_0; T')$ is a space with involution.

5.5. The covering map p induces functions:

$$p_\#\colon \pi_n{}^i(Y, y_0; T') \to \pi_n{}^i(X, x_0; T) \qquad \text{if } i \geqq 1,$$
$$p_\#\colon \pi_n{}^0(Y, \bar{T}) \to \pi_n{}^0(X; T).$$

In the spirit of homotopy theory, we may say a great deal about $p_\#$.

PROPOSITION. A. (i) *If $i \geqq 1$, $p_\#$ is a monomorphism.*
(ii) *If $i \geqq 2$, $p_\#$ is an epimorphism.*
B. (i) *If $i = 0$, $p_\#(\beta) = p_\#(\beta')$ if and only if $\beta = \beta'$ or $\beta = t_\#(\beta')$.*
(ii) *If $i = 0$ and $n \geqq 2$, $\pi_n{}^0(X; T)$ is the union of the images under $p_\#$ of $\pi_n{}^0(Y; T')$ and $\pi_n{}^0(Y; T'')$.*

Proof. We first prove A-(ii) and B-(ii). Let $f \in F_n{}^i(X, x_0; T)$, where $n \geqq 2$ and $i \neq 1$. Since S^n is simply-connected and S^{i-1} is connected, [3, p. 89] says that there is a map $g\colon (S^n, S^{i-1}) \to (Y, y_0)$ such that $pg = f$. If $i > 0$, S^{i-1} is non-empty and [3, p. 89] says g is unique. But $T'gT_1$ is such a map also; therefore $g \in F_n{}^i(Y, y_0; T')$ and A-(ii) is proven. If $i = 0$, [3, p. 89] says there are exactly *two* mappings with the properties of g. But $T'gT_0 \neq T''gT_0$; therefore $g \in F_n{}^0(Y, \bar{T})$ for some $\bar{T}$, and B-(ii) is proven.

We now prove A-(i) and B-(i). Let $g_0, g_1 \in F_n{}^i(Y, y_0; \bar{T})$, where $\bar{T} = T'$ if $i > 0$. Suppose there is a homotopy $\{f_t\} \in F_n{}^i(X, x_0; T)$ between $f_0 = pg_0$

and $f_1 = pg_1$. By [3, p. 89], there is a unique homotopy $g'_t\colon (S^n, S^{i-1}) \to (Y, y_0)$ with $g'_0 = g_0$. But $\{Tg'_tT_i\}$ is also such a homotopy; therefore $\{g'_t\} \in F_n{}^i(Y, y_0; T)$. Now $pg'_1 = pg_1$ and $g_1(S^{i-1}) = g'_1(S^{i-1})$. If $i > 0$, [3, p. 89] says $g'_1 = g_1$; if $i = 0$, $g'_1 = g_1$ or tg_1. This completes the proof.

Part III.

6. The orthogonal group.

6.1. The remainder of the paper will be devoted to calculating $\pi_{n-1}{}^0(X;T)$ where X is a component of the orthogonal group and T matrix inversion. By 2.6, the operation Θ converts this into a classification of vector bundles over P^n. Complete computations will be made for $n \leqq 4$.

In this section we will study certain structure induced on the $\pi_{n-1}{}^0(X;T)$ by Θ.

6.2. Let O_k denote the orthogonal group on k variables. We denote by R_k and A_k, the two components of O_k, consisting of rotations and reflections, respectively; X_k will denote either of the components. Matrix inversion, denoted by T, induces on involution of X_k. It is clear that $\pi_n{}^0(O_k)$, in the notation of 2.6, is the union of the $\pi_n{}^0(X_k;T)$. We shall write $\pi_n{}^i(X_k) = \pi_n{}^i(X_k, x_0; T)$.

Since the orientable k-plane bundles are those with group R_k, $\Theta(\pi_{n-1}{}^0(R_k))$ *are the orientable bundles* and, consequently, $\Theta(\pi_{n-1}{}^0(A_k))$ *are the non-orientable bundles.*

6.3. Let $\alpha \in \pi_n{}^0(X_k)$, $\beta \in \pi_n{}^0(X_m)$ be represented by $f \in F_n{}^0(X_k)$ and $g \in F_n{}^0(X_m)$, respectively. Consider $h \in F_n{}^0(X_{k+m})$, defined by:

$$h(x) = \begin{pmatrix} f(x) & 0 \\ 0 & g(x) \end{pmatrix}$$

If h represents $\gamma \in \pi_n{}^0(X_{k+m})$, we write $\gamma = \alpha \oplus \beta$. This gives an associative and commutative pairing.

PROPOSITION. $\Theta(\alpha \oplus \beta)$ *is the Whitney sum of* $\Theta(\alpha)$ *and* $\Theta(\beta)$.

Proof. This is a direct consequence of the definition of Whitney sum as follows. If two bundles are given by coordinate neighborhoods indexed by a set I and coordinate transformations $\{g'_{ij} \mid i, j \in I\}$ and $\{g''_{ij} \mid i, j \in I\}$, their Whitney sum is given by coordinate transformations $\{g_{ij} \mid i, j \in I\}$ where:

$$g_{ij}(x) = \begin{pmatrix} g'_{ij}(x) & 0 \\ 0 & g''_{ij}(x) \end{pmatrix}$$

Let e_m denote the identity matrix in O_m; as in 3.3, e_m also denotes elements of $F_n{}^i(R_m; e_m)$ and $\pi_n{}^i(R_m, e_m)$. If $i\colon O_k \to O_{k+m}$ is the natural inclusion, then:

$$i_{\#}\colon \pi_n{}^0(X_k) \to \pi_n{}^0(X_{k+m})$$

is defined and, it is easy to see, $i_{\#}(\alpha) = \alpha \oplus e_m$, for $\alpha \in \pi_n{}^0(X_k)$.

6.4. We now discuss Stiefel-Whitney classes ([4]). Let $L_n[u]$ be the polynomial ring over Z_2 on the indeterminant u and truncated by $u^{n+2} = 0$. The Stiefel-Whitney class of a k-plane bundle over P^{n+1} can be considered an element of $L_n[u]$ if u is identified with the generator of $H^{*}(P^{n+1}; Z_2)$. *We define:*

$$W\colon \pi_n{}^0(X_k) \to L_n[u]$$

by $W(\alpha) =$ *Stiefel-Whitney class of* $\Theta(\alpha)$.

The following propositions are restatements of well-known facts in our terminology.

PROPOSITION 1. *The following diagram is commutative:*

$$\begin{array}{ccc}
\pi_n{}^0(X_k) & \xrightarrow{\psi_0} & \pi_{n-1}{}^0(X_k) \\
\downarrow W & & \downarrow W \\
L_n[u] & \xrightarrow{\psi} & L_{n-1}[u]
\end{array}$$

where ψ *is the reduction* $u^{n+1} = 0$.

PROPOSITION 2. *If* $\alpha \in \pi_n{}^0(X_k)$, $\beta \in \pi_n{}^0(X_m)$, *then*

$$W(\alpha \oplus \beta) = W(\alpha) \cdot W(\beta) \text{ in } L_n[u].$$

Proposition 1 is obvious if we notice that ψ_0 corresponds, under Θ, to the restriction of bundles over P^{n+1} to P^n. Proposition 2 is the Whitney product theorem.

6.5. Let $f \in F_n{}^0(O_k)$. If we define g by:

$$g(x) = (-e_k) \cdot f(x) = f(x) \cdot (-e_k)$$

using matrix multiplication, clearly $g \in F_n{}^0(O_k)$. Thus $f \to g$ defines an involution I of $\pi_n{}^0(O_k)$.

PROPOSITION. *If k is odd, I induces a one-one correspondence between* $\pi_n{}^0(R_k)$ *and* $\pi_n{}^0(A_k)$.

Proof. If k is odd, $-e_k \in A_k$ and I interchanges the components of O_k.

6.6. The following "stability" theorem follows directly from

5.2. THEOREM. *Let $i\colon O_k \to O_{k+1}$ be the natural inclusion and $a \in X_k$ any element satisfying $a^2 = e_k$. Then:*

$$i_\#\colon \pi_n{}^i(X_k, a) \to \pi_n{}^i(X_{k+1}, a)$$

(a) *is onto if $k \geqq n+1$.*

(b) *is one-one if $k \geqq n+2$.*

6.7. If $k \geqq 1$, let a_k denote the element $-e_k$ in O_k. If $m \geqq k$, let a_k denote the element $i(-e_k)$ in O_m, where i is the natural inclusion $O_k \subset O_m$. As in 3.3, a_k also denotes elements of $F_n{}^i(X_m, a_k)$ and $\pi_n{}^i(X_m, a_k)$, since $a_k{}^2 = e_m$.

Let L be the non-trivial line bundle over P^{n+1}.

PROPOSITION. *Let $a_k \in \pi_n{}^0(X_m)$; then $\otimes(a_k)$ is the Whitney sum of k copies of L and a trivial bundle and $\otimes(e_m)$ is the trivial bundle. Consequently, $W(e_m) = 1$ and $W(a_k) = (1+u)^k$.*

Proof. If $k = m = 1$, $R_m = \{e_1\}$ and $A_m = \{a_1\}$ and the proposition is obvious. For $m \geqq k \geqq 1$, the proposition follows from 6.3 and 6.4 if we observe

$$a_k = a_1 \oplus \cdots \oplus a_1 \oplus e_1 \oplus \cdots \oplus e_1$$

and $e_m = e_1 \oplus \cdots \oplus e_1$.

where we have m components in each decomposition and k copies of a_1.

7. Computations.

7.1. We will now compute the $\pi_n{}^0(X_k)$ for $n \leqq 3$. If S is a set, we write:

$$S = \{C_1, C_2, \cdots, C_r, \cdots\}$$

to indicate that S is composed of the *distinct* elements $C_1, C_2, \cdots, C_r, \cdots$. In presenting our results we will use the notation of the group action introduced in 3.8. We will also need to fix notation for elements of $\pi_n(X_k)$.

(a) $\pi_1(A_2) = Z$ generated by ϵ.

(b) $\pi_3(R_3) = Z$ generated by σ_0.

(c) $\pi_3(R_4) = Z + Z$ generated by σ_1 and σ_2.

(d) $\pi_3(R_k) = Z$ generated by σ, for $k \geqq 5$.

In addition, under the inclusion $R_3 \subset R_4$, $\sigma_0 \to \sigma_1$; under $R_4 \subset R_5$, $\sigma_2 \to \sigma$ and $\sigma_1 \to -2\sigma$. We refer the reader to [6] for proofs and more detail. In all cases, if $\alpha \in \pi_n(R_k)$, then α' will denote the translate of α into $\pi_n(A_k)$.

A. 1. $\pi_n{}^0(R_2) = \{e_2, a_2\}$ for $n \geqq 1$.

2. (a) $\pi_1{}^0(A_2) = \{(n\epsilon) \cdot a_1 \mid n \in Z\}$
(b) $\pi_n{}^0(A_2) = \{a_1\}$ for $n \geqq 2$.

3. Action of $\pi_0(O_2)$ is trivial except on $\pi_1{}^0(A_2)$ where $(n\epsilon) \cdot a_1 \to (-n\epsilon) \cdot a_1$.

B. 1. $\pi_n{}^0(R_k) = \{e_k, a_2\}$ for $k \geqq 2$, $n = 1, 2$.
2. $\pi_n{}^0(A_k) = \{a_1, a_3\}$ for $k \geqq 3$, $n = 1, 2$.
3. Action of $\pi_0(O_k)$ is trivial.

C. 1. (a) $\pi_3{}^0(R_3) = \{e_3, a_2, \sigma_0 \cdot a_2\}$
(b) $\pi_3{}^0(R_4) = \{e_4, a_2, \sigma_1 \cdot a_2, \sigma_2 \cdot a_2, (\sigma_1 + \sigma_2) \cdot a_2, a_4\}$
(c) $\pi_3{}^0(R_k) = \{e_k, a_2, \sigma \cdot a_2, a_4\}$ for $k \geqq 5$.
and $\sigma \cdot a_2 = a_6$ if $k \geqq 6$.

2. (a) $\pi_3{}^0(A_3) = \{a_1, a_3, \sigma'_0 \cdot a_1\}$
(b) $\pi_3{}^0(A_4) = \{(n\sigma'_3) \cdot a_1, (n\sigma'_2) \cdot a_3 \mid n \in Z\}$
(c) $\pi_3{}^0(A_k) = \{a_1, a_3, \sigma' \cdot a_3, a_5\}$ for $k \geqq 5$.
and $\sigma' \cdot a_3 = a_7$ if $k \geqq 7$.

3. Action of $\pi_0(O_k)$ is trivial except on $\pi_3{}^0(R_4)$ where $\sigma_2 \cdot a_2 \to (\sigma_1 + \sigma_2) \cdot a_2$ and on $\pi_3{}^0(A_4)$ where $(n\sigma'_2) \cdot a_k \to (-n\sigma'_2) \cdot a_k$.

There are many more details which may be easily deduced from the proofs to be presented e. g. the Stiefel-Whitney and Euler classes and the effect of the natural inclusions $O_k \subset O_{k+1}$. In 7.7, we present a restatement of our results after applying $\otimes$, i. e. in the language of vector bundles.

It is also quite easy to calculate $\pi_n{}^0(X_k)$ for $n \leqq 7$, in the stable range $k \geqq n + 2$, but these are already known (see [1]).

7.2. We introduce a notational convention. Because of the ambiguity in our notation for the functions ϕ_i, ψ_i and ρ_i, we may be led to discussing, in the same context, more than one of these functions with the same name. In this case we will write one as $\bar{\phi}_i$, $\bar{\psi}_i$ or $\bar{\rho}_i$ to differentiate them.

Proof of A. Let $a \in X_2$ such that $a^2 = e_2$. Consider the following diagram:

$$\begin{tikzcd}
\pi_1{}^1(X_2, a) \arrow[r, "\rho_1"] & \pi_1(X_2) \arrow[r, "\phi_0"] & \pi_1{}^0(X_2, a) \arrow[r, "\psi_0"] & \pi_0{}^0(X_2) \\
\pi_1(X_2) \arrow[u, "\phi_1"] & & & \pi_0(X_2) \arrow[u, "\bar{\phi}_0"']
\end{tikzcd}$$

By 4.3, the row is exact; by (4.8-b), $\bar{\phi}_0$ and ϕ_1 are isomorphisms. Then exactness tells us ϕ_0 is onto i.e. $\pi_1{}^0(X_2)$ is entirely the orbit of a under the action of $\pi_1(X_2)$.

Since matrix inversion is of degree -1 on R_2 and the identity on A_2 we have, from (4.8-a):

$$\rho_1\phi_1(\alpha) = \begin{cases} 0 & \text{if } \alpha \in \pi_1(A_2), \\ 2\alpha & \text{if } \alpha \in \pi_1(R_2). \end{cases}$$

Since ϕ_1 is an isomorphism, this determines Image ρ_1. By exactness, this proves 2(a). If $X_2 = R_2$, it shows $\pi_1{}^0(R_2)$ consists of *two* elements; but $W(e_2) \neq W(a_2)$, by 6.7. This verifies 1 for $n = 1$.

Now consider the following exact sequence:

$$\pi_n(X_2) \xrightarrow{\phi_0} \pi_n{}^0(X_2, a) \xrightarrow{\psi_0} \pi_{n-1}{}^0(X_2) \xrightarrow{\rho_0} \pi_{n-1}(X_2).$$

Note that $\rho_0(e_k) = \rho_0(a_k) = 0$. Let $n = 2$; then if $\alpha \in \pi_1(A_2)$, $\rho_0(\alpha \cdot a_1) = 2\alpha$ by (4.8-a). It then follows that:

$$\text{Image}\,\psi_0 = \text{Ker}\,\rho_0 = \begin{cases} \{e, a_2\} & \text{if } X_2 = R_2, \\ \{a_1\} & \text{if } X_2 = A_2. \end{cases}$$

Since $\pi_2(X_2) = 0$, ψ_0 is a monomorphism and 1 and 2 follow for $n = 2$. If $n \geqq 3$, $\pi_n(X_2) = \pi_{n-1}(X_2) = 0$ and ψ_0 is an isomorphism; then 1 and 2 follow for $n \geqq 3$.

To verify 3, we use the following two easily verified facts.

(i) $\pi_0(O_m)$ acts trivially on e_m and a_k, for $k \leqq m$.

(ii) If $\alpha \in \pi_n(X_k)$, $\beta \in \pi_n{}^0(X_k)$ and the action of $\pi_0(O_k)$ gives $\alpha \to \alpha_0$, $\beta \to \beta_0$, then $\alpha \cdot \beta \to \alpha_0 \cdot \beta_0$.

Since all the elements in the $\pi_n{}^0(X_k)$ which will be investigated will be of the form $\alpha \cdot e_m$ or $\alpha \cdot a_k$, statement 3 of A, B and C follows directly from (i) and (ii) and knowledge of the action of $\pi_0(O_k)$ on $\pi_n(X_k)$.

For $n=1$, consider the following diagram for $k \geqq 3$, with exact row:

$$\begin{array}{ccccc} \pi_1(X_k) & \xrightarrow{\phi_0} & \pi_1{}^0(X_k) & \xrightarrow{\psi_0} & \pi_0{}^0(X_k) \\ & & & & \uparrow \bar{\phi}_0 \\ & & & & \pi_0(X_k) \end{array}$$

Since $\bar{\phi}_0$ is an isomorphism by (4.8-b) and $\pi_1(X_k) \approx Z_2$, $\pi_1(X_k)$ consists of at most *two* elements. But $w(e_2) \neq w(a_2)$ and $w(a_1) \neq w(a_3)$, by 6.7; this verifies B for $n=1$.

7.3. *Proof of B.* For $n=2$ consider the following exact sequence for $k \geqq 3$:

$$\pi_2(X_k) \xrightarrow{\phi_0} \pi_2{}^0(X_k, a) \xrightarrow{\psi_0} \pi_1{}^0(X_k) \xrightarrow{\rho_0} \pi_1(X_k).$$

We have just seen that $\pi_1{}^0(X_k)$ consists of "constant" elements; therefore $\rho_0 = 0$. Since $\pi_2(X_k) = 0$, ψ_0 is an isomorphism and B is proved.

7.4. *Proof of* C. To prove 1(a) consider the following exact sequence:

$$\pi_3{}^1(R_3, a) \xrightarrow{\rho_1} \pi_3(R_3) \xrightarrow{\phi_0} \pi_3{}^0(R_3, a) \xrightarrow{\psi_0} \pi_2{}^0(R_3) \xrightarrow{\rho_0} \pi_2(R_3).$$

Since $\pi_2{}^0(R_3)$ consists of "constants," $\rho_0 = 0$; by exactness, $\pi_3{}^0(R_3)$ consists on the orbits of e_3 and a_2. By (4.8-a), $\rho_1(\alpha \cdot a) = 2\alpha$ for $\alpha \in \pi_3(R_3)$, $a = e_3$ or a_2; thus the orbit of a is a and $\sigma_0 \cdot a$. To complete the proof of 1(a) we must show:

(i) $e_3 = \sigma_0 \cdot e_3$.

(ii) $a_2 \neq \sigma_0 \cdot a_2$.

Proof of (i). By exactness, we must show $\sigma_0 \in \text{Image}\, \rho_1$ for $a = e_3$ in the above sequence. Now, according to [6], $\bar{\rho}$ is represented by the following map $f\colon S^3 \to R_3$. Identify S^3 with the unit quaternions, where the *first* coordinate is the *reals*, and R_3 *with* the rotations of the *purely imaginary* quaternions. Then f is defined by:

$$f(q) \cdot q' = qq'\bar{q}$$

where q, q' are unit quaternions and $\bar{q}$ is the conjugate and inverse of q. Since $T_1(q) = \bar{q}$ and $S^0 = \{\pm 1\}$, it is straightforward to check that $f \in F_3{}^1(R_3, e\ )$. But this means $\sigma_0 \in \text{Image}\, \rho_1$, which proves (i).

Proof of (ii). Consider the following diagram with exact row:

$$\begin{array}{ccccc}
\pi_4{}^0(R_3) & \xrightarrow{\psi_0} & \pi_3{}^0(R_3) & \xrightarrow{\rho_0} & \pi_3(R_3) \\
\Big\downarrow{\scriptstyle \bar{\rho}_0} & & & & \\
\pi_4(R_3) & & & &
\end{array}$$

Suppose $a_2 = \sigma_0 \cdot a_2$; then $\pi_3{}^0(R_3) = \{e_3, a_2\}$ and $\rho_0 = 0$. By exactness, $\pi_4{}^0(R_3)$ consists of the orbits of e_3 and a_2. It is then easy to see from (4.8-a) that $\bar{\rho}_0 = 0$, since $\pi_4(R_3) = Z_2$. We will contradict this by exhibiting an essential map $g\colon S^4 \to R_3$ such that $g \in F_4{}^0(R_3)$.

Let $h\colon S^3 \to S^2$ be the Hopf map which collapses great circles; in particular $hT_0 = h$. Let $h'\colon S^4 \to S^3$ be the *suspension* of h along the first coordinate; because it is a suspension, $h'(T_1T_0) = (T_0T_1)h'$. From $hT_0 = h$, it follows that $h'T_1 = h'$, since T_1 is the suspension of T_0. By composing with T_0, we have:

$$(*) \qquad h'T_0 = T_0T_1h'$$

Now recall the map f from the proof fo (i); it is easy to check that $fT_0 = f$. By applying this to (*) we have:

$$(fh')T_0 = fT_1h' = T(fh').$$

Thus $fh' \in F_4{}^0(R_3)$; but, according to [6], fh' is an essential map. This proves (ii) and, thus, 1(a).

7.5. To prove 1(b), we will study the covering $p\colon S^3 \times S^3 \to R_4$. If we identify S^3 with the unit quarternions, as in 7.4 and R_4 with the rotations of the unit quaternions:

$$p(q_1, q_2) \cdot q = q_1 q \bar{q}_2 \qquad \text{for } q, q_1, q_2 \in S^3.$$

The covering transformation $t\colon S^3 \times S^3 \to S^3 \times S^3$ is defined by:

$$t(q_1, q_2) = (-q_1, -q_2) \qquad \text{for } q_1, q_2 \in S^3.$$

We can apply the theory of 5.4 and 5.5. It is easy to check that the involutions T' and T'' covering T are defined by:

$$\begin{aligned} T'(q_1, q_2) &= (\bar{q}_1, \bar{q}_2) \\ T''(q_1, q_2) &= (-\bar{q}_1, -\bar{q}_2) \qquad \text{for } q_1, q_2 \in S^3. \end{aligned}$$

Let $\bar{T}$ denote either T' or T'' and let $X = S^3 \times S^3$.

LEMMA 1. *If b is a fixed point of $\bar{T}$ and α_1, α_2 generate $\pi_3(X)$,*

$$\pi_3{}^0(X;\bar{T}) = \{b, \alpha_1 \cdot b, \alpha_2 \cdot b, (\alpha_1 + \alpha_2) \cdot b\}.$$

Proof. Since X is 2-connected, it follows easily from 4.3 and (4.8-b) that:

$$(\dagger) \qquad \pi_n{}^i(X, b; \bar{T}) = \{b\} \qquad \text{if } i \leqq n \leqq 2.$$

Now consider the following diagram with exact row and column:

$$\begin{array}{lllllll}
\pi_3(X) & & & & & & \\
\big\downarrow \phi_1 & & & & & & \\
\pi_3{}^1(X, b; \bar{T}) & \xrightarrow{\rho_1} & \pi_3(X) & \xrightarrow{\phi_0} & \pi_3{}^0(X, b; \bar{T}) & \xrightarrow{\psi_0} & \pi_2{}^0(X; \bar{T}) \\
\big\downarrow \psi_1 & & & & & & \\
\pi_3{}^1(X, b; \bar{T}). & & & & & &
\end{array}$$

From $(\dagger)$, ϕ_1 and ϕ_0 are onto; by (4.8-a), $\rho_1\phi_1(\alpha) = 2\alpha$. The lemma follows immediately.

Let $b', b'' \in X$ be fixed points of T', T'', respectively, so that $p(b') = e_4$ and $p(b'') = a_2$. Choose α_1, α_2 so that $p_*(\alpha_j) = \sigma_j$. Let t be the covering transformation defined above.

LEMMA 2. (i) $t_\#(b') = \alpha_1 \cdot b'$; (ii) $t_\#(b'') = b''$.

Proof. (i) Since $p_\#(\alpha_1 \cdot b') = \sigma_1 \cdot e_4$ and $p_\#(b') = e_4$, it suffices, by (5.5B-(i)), since $\alpha_1 \cdot b' \neq b'$ by Lemma 1, to show $\sigma_1 \cdot e_4 = e_4$. But this follows from (i) of 7.4 under the inclusion $R_3 \subset R_4$.

(ii) The fixed point set of T'' is homeomorphic to $S^2 \times S^2$; in particular, it is *connected*. A path in this set between b'' and $t(b'')$ defines a homotopy of constant maps in $F_3{}^0(X; T'')$ which proves (ii).

Since t_* on $\pi_3(X)$ is the identity, Lemmas 1 and 2 and the formula of 5.1 describe $t_\#$ completely. Together with (5.5-B) this implies:

$$\pi_3{}^0(R_4) = \{e_4, \sigma_2 \cdot e_4, a_2, \sigma_1 \cdot a_2, \sigma_2 \cdot a_2, (\sigma_1 + \sigma_2) \cdot a_2\}.$$

To complete the proof of 1(b) we have to show $\sigma_2 \cdot e_4 = a_4$. Consider the following exact sequence:

$$\pi_3(R_4) \xrightarrow{\phi_0} \pi_3{}^0(R_4, e_4) \xrightarrow{\psi_0} \pi_2{}^0(R_4).$$

By definition, $\sigma_2 \cdot e_4 = \phi_0(\sigma_2)$; therefore $\psi_0(\sigma_2 \cdot e_4) = e_4$ and, by 6.4, $\psi W(\sigma_2 \cdot e_4)$

$=1$. Similarly $\psi W(\alpha\cdot a_2)=1+u^2$, for any $\alpha\in\pi_3(R_4)$. But $\psi W(a_4)=\psi(1+u^4)=1$; since $W(e_4)\neq W(a_4)$, this shows $a_4=\sigma_2\cdot e_4$.

To prove 1(c) we merely apply 6.6 and 7.1 to 1(b) and note that $\pi_3{}^0(R_k)$ has at least four distinct elements, since $W(e_k)$, $W(a_2)$, $W(a_4)$ and $W(a_0)$ are distinct.

7.6. We now prove 2 of C. In fact, (a) and (c) follow from 1(a) and (c) and 6.5. It remains only to prove (b).

Consider the following diagram with exact row where $a=a_1$ or a_3:

$$\begin{array}{ccccccc}\pi_3{}^1(A_4,a) & \xrightarrow{\rho_1} & \pi_3(A_4) & \xrightarrow{\phi_0} & \pi_3{}^0(A_4,a) & \xrightarrow{\psi_0} & \pi_2{}^0(A_4)\\ & & & & \Big\downarrow{\scriptstyle \rho_0} & & \\ & & & & \pi_3(A_4). & & \end{array}$$

By (B, 2), ψ_0 is onto and $\pi_3{}^0(A_4)$ consists of the orbits of a_1 and a_3. Clearly 2(b) will follow from the following two statements.

(i) $\sigma'_1\in\operatorname{Image}\rho_1$ for $a=a_1$ and a_3.

(ii) $\rho_0\phi_0(\sigma'_2)\neq 0$ for $a=a_1$ and a_3.

Proof of (i). If $a\in A_4$, define $f_a\colon S^3\to A_4$ by:

$$f_a(q)=a\cdot f(q)\qquad\text{for } q\in S^3,$$

where f is defined in 7.4. If $a=a_1$ or $-a_1$, $f_a\in F_3{}^1(A_4,a)$ since $a_1(q)=-\bar q$ and $(-a_1)(q)=\bar q$. For $a=a_1$, this proves (i). Let A be a rotation which satisfies $Aa_3A^{-1}=-a_1$; e.g. a cyclic permutation of the first, third and fourth coordinates. Now define $g\colon S^3\to A_4$ by $g=A^{-1}f_aA$, where $a=-a_1$. It is easy to check that $g\in F_3{}^1(A_4,a_3)$, which proves (i) for a_3.

Proof of (ii). Since $T_*(\sigma'_2)=\sigma'_2$, it follows from (4.8-a) that $\rho_0\phi_0(\sigma'_2)=2\sigma'_2$.

This completes the proof of C.

7.7. We conclude this work with a tabulation of the results of 7.1 in terms of vector bundles over P^n. T and L denote the trivial and non-trivial line bundles, resp., and summation is Whitney sum. The $\{A_i\}$, $\{B_i\}$ and $\{C_i\}$ are infinite families of vector bundles; the subscript i is the *Euler class* of the bundle, assuming an identification of the suitable cohomology group with the integers. All other bundles are denoted by the symbol $\otimes$; those with asterisks are not distinguishable by their stable class or any characteristic

class. Finally, for a given value of n, two bundles appear in the *same row* if and only if they are stably equivalent.

k-plane bundles over P^n for $n \leqq 4$.

$n = 2$

$k = 2$	$k \geqq 3$
$2T$	kT
$L + T, \{A_{2i} \mid i > 0\}$	$L + (k-1)T$
$2L$	$2L + (k-2)T$
$\{A_{2i-1} \mid i > 0\}$	$3L + (k-3)T$

$n = 3$

$k = 2$	$k \geqq 3$
$2T$	kT
$L + T$	$L + (k-1)T$
$2L$	$2L + (k-2)T$
	$3L + (k-3)T$

$n = 4$

$k = 2$	$k = 3$	$k = 4$	$k \geqq 5$
$2T$	$3T$	$4T$	kT
$L + T$	$L + 2T, \otimes^{(*)}$	$L + 3T, \{B_{2i} \mid i > 0\}$	$L + (k-1)T$
$2L$	$2L + T, \otimes^{(*)}$	$2L + 2T, \otimes^{(*)}$	$2L + (k-2)T$
	$3L$	$3L + T, \{C_{2i} \mid i > 0\}$	$3L + (k-3)T$
		$4L$	$4L + (k-4)T$
		$\{B_{2i-1} \mid i > 0\}$	$5L + (k-5)T$
		$\otimes$	$\otimes \rightarrow 6L + (k-6)T$
		$\{C_{2i-1} \mid i > 0\}$	$\otimes \rightarrow 7L + (k-7)T$

Remark. The Euler classes are calculated by examining:

$$\pi_{n-1}{}^0(A_n) \xrightarrow{\rho_0} \pi_{n-1}(A_n) \xrightarrow{p_*} \pi_{n-1}(S^{n-1})$$

where p is the usual projection, and using the fact that ρ_0 corresponds, under Θ, to passing to the bundle over S^n induced by the two-fold covering $S^n \rightarrow P^n$.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] J. F. Adams, "Vector fields on spheres," *Annals of Mathematics*, vol. 75 (1962), pp. 603-632.

[2] A. M. Gleason, "Spaces with a compact Lie group of transformations," *Proceedings of the American Mathematical Society*, vol. 1 (1950), pp. 35-43.

[3] S. T. Hu, *Homotopy theory*, Academic Press, 1959.

[4] J. Milnor, *Lectures on characteristic classes*, mimeographed notes, Princeton University, 1957.

[5] D. Puppe, "Homotopiemengen und ihre induzierten Abbildungen I," *Mathematische Zeitschrift*, vol. 69 (1958), pp. 299-344.

[6] N. Steenrod, *Topology of fibre bundles*, Princeton University Press, 1951.

A FINITELY PRESENTED GROUP WHOSE 3-DIMENSIONAL INTEGRAL HOMOLOGY IS NOT FINITELY GENERATED.*

By John Stallings.[1]

The aim of this note is to provide a counterexample to a conjecture about the niceness of finitely presented groups. This also gives a counterexample to a conjecture about $\pi_2(K)$, where K is a finite complex.

Example. *The group G with presentation*

$$\{a, b, c, x, y : [x, a], [y, a], [x, b], [y, b], [a^{-1}x, c], [a^{-1}y, c], [b^{-1}a, c]\}$$

with five generators and seven relations has as its 3-dimensional homology group with integer coefficients a not finitely generated group. (Note: $[u, v] = uvu^{-1}v^{-1}$.)

Corollary 1. *There is no projective resolution* [1] *of Z over $Z(G)$ which is finitely generated in dimension* 3.

Corollary 2. *If K is any finite complex with $\pi_1(K) \approx G$, then $\pi_2(K)$ is not finitely generated, even as a module over $\pi_1(K)$.*

1. The Mayer-Vietoris sequence of a free product with Amalgamation. The notation $(A * B)_C$ denotes the free product of the groups A and B with amalgamated subgroup C ([4], p. 29). $\mathrm{Dub}(A, C)$ denotes $(A_1 * A_2)_C$ where A_1 and A_2 are two copies of A each containing the subgroup C.

For any group A there is a cell complex $\mathfrak{A}$ such that $\pi_1(\mathfrak{A}) \approx A$, $\pi_n(\mathfrak{A}) = 0$ for $n \neq 1$; let $\mathfrak{B}$ and $\mathfrak{C}$ be complexes with similar properties with respect to the groups B and C. Let $f : \mathfrak{C} \to \mathfrak{A}$ and $g : \mathfrak{C} \to \mathfrak{B}$ be cellular maps inducing certain monomorphisms $C \to A$, $C \to B$. Let $\mathfrak{D}$ denote the union of the mapping cylinders $\mathrm{Map}(f)$, $\mathrm{Map}(g)$, identified along their isimorphic subcomplexes which are copies of $\mathfrak{C}$. There is a Mayer-Vietoris sequence ([2], p. 37) for the triad $(\mathfrak{D}; \mathrm{Map}(f), \mathrm{Map}(g))$.

* Received October 12, 1961.

[1] This work was supported by a National Science Foundation fellowship.

$$\cdots \to H_{n+1}(\mathfrak{D}) \to H_n(\mathfrak{C}) \to H_n(\mathrm{Map}(f)) \oplus H_n(\mathrm{Map}\,(g)) \to \cdots.$$

The coefficient group is understood to be a trivial module over all the fundamental groups involved.

Now,

$$H_n(\mathfrak{C}) \approx H_n(C); \qquad H_n(\mathrm{Map}(f)) \approx H_n(\mathfrak{A}) \approx H_n(A);$$
$$H_n(\mathrm{Map}(g)) \approx H_n(\mathfrak{B}) \approx H_n(B).$$

Finally, since $\mathfrak{D}$ is aspherical ([5], p. 160), and since $\pi_1(\mathfrak{D}) \approx (A * B)_C$, $H_n(\mathfrak{D}) \approx H_n((A * B)_C)$. Thus we get the following Mayer-Vietoris sequence:

$$\cdots \to H_{n+1}((A * B)_C) \to H_n(C) \to H_n(A) \oplus H_n(B) \to \cdots.$$

2. Building up G. Let A denote the free group with basis $\{a, b\}$. And let B denote its subgroup generated by $\{a^n b a^{-n}, n = 0, \pm 1, \pm 2, \cdots\}$; B is freely generated by these elements and hence has infinite rank. Let C denote $\mathrm{Dub}(A, B)$; note the following part of the Mayer-Vietoris sequence, with integer coefficients:

$$H_2(C) \to H_1(B) \to H_1(A) \oplus H_1(A)$$

Since $H_1(B)$ is not finitely generated and $H_1(A)$ is finitely generated, it follows that $H_2(C)$ *is not finitely generated.* The following is a presentation of C:

$$\{a_1, a_2, b : a_1{}^n b a_1{}^{-n} = a_2{}^n b a_2{}^{-n}, n = 0, \pm 1, \pm 2, \cdots\}.$$

Define a group D by the presentation:

$$\{a_1, a_2, b, x : x a_1 x^{-1} = a_1, x a_2 x^{-1} = a_2, x b x^{-1} = a_1 b a_1{}^{-1}, a_1 b a_1{}^{-1} = a_2 b a_2{}^{-1}\}.$$

The subgroup generated by $\{a_1, a_2, b\}$ in D can be easily shown to be the group C, by the standard method of writing down a presentation of a subgroup ([3]). Since D is finitely presented, $H_2(D)$ is finitely generated. Define $G = \mathrm{Dub}(D, C)$. Note the following part of the MayerVietoris sequence for G.

$$H_3(G) \to H_2(C) \to H_2(D) \oplus H_2(D).$$

$H_2(C)$ is not finitely generated, while $H_2(D)$ is finitely generated. Hence $H_3(G)$ *is not finitely generated.* A presentation of G is:

$$\{a_1, a_2, b, x_1, x_2 : x_1a_1x_1^{-1} = a_1, x_2a_1x_2^{-1} = a_1, x_1a_2x_1^{-1} = a_2,$$
$$x_2a_2x_2^{-1} = a_2, x_1bx_1^{-1} = a_1ba_1^{-1}, x_2bx_2^{-1} = a_1ba_1^{-1}, a_1ba_1^{-1} = a_2ba_2^{-1}\}.$$

The presentation given at the beginning of this note is a modified form of this.

The first corollary is an evident consequence of the property of $H_3(G)$. In the second corollary, if $\pi_2(K)$ were finitely generated over $\pi_1(K)$, then by adding a finite number of 3-cells to K, and some n-cells for $n > 3$, one would obtain an aspherical complex $\mathfrak{G}$ with finite 3-skeleton and $\pi_1(\mathfrak{G}) \approx G$; and this would imply that $H_3(G)$ is finitely generated.

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton, 1956.

[2] S. Eilenberg and N. Steenrod, *Foundations of algebraic topology*, Princeton, 1956.

[3] R. H. Fox, "Free differential calculus III. subgroups," *Annals of Mathematics*, vol. 64 (1956), pp. 407-419.

[4] A. G. Kurosh, *The theory of groups*, Chelsea, 1956. Vol. 2.

[5] J. H. C. Whitehead, "On the asphericity of regions in the 3-sphere," *Fundamental Mathematics*, vol. 32 (1939).

ON A CERTAIN SUBGROUP OF THE GROUP OF AN ALTERNATING LINK.*

By KUNIO MURASUGI.[1]

1. **Introduction.** Let $L \subset S^3$ be a tame oriented link of multiplicity μ. Let $G = (x_1, x_2, \cdots, x_n : r_1, r_2, \cdots, r_m)$ be a *Wirtinger* presentation of the group $G = \pi_1(S^3 - L)$ of L, and let θ be the homomorphism of G onto an infinite cyclic group $Z = (t:)$ defined by $\theta(x_i) = t$ for all i. Let H be the kernel of θ. In other words, H is the (normal) subgroup of G, an element of which is represented by a loop l in $S^3 - L$ such that the sum of the linking numbers of l with each component of L is zero. Clearly θ can be uniquely extended to the homomorphism $\tilde{\theta}$: $JG \to JZ$ of the integral group rings. Then, denoting the Jacobian matrix of $G = (x:r)$ by $\| \partial r_i / \partial x_j \|$, the g. c. d. of the determinants of all $(n-1) \times (n-1)$ minor of $\| \tilde{\theta}(\partial r_i / \partial x_j) \|$ is called the *reduced Alexander polynomial* of L, denoted by $\tilde{\Delta}(t)$. Cf. [1], [3]. Since $\tilde{\Delta}(t)$ is defined only to within an arbitrary factor $\pm t^{\lambda}$, we may assume that $\tilde{\Delta}(0) \neq 0$, unless $\tilde{\Delta}(t) = 0$.

The object of this paper is to show the following

THEOREM 1.1. *Suppose L is an alternating link of multiplicity μ, and let d be the degree of its polynomial $\tilde{\Delta}(t)$. If $\tilde{\Delta}(0) = \pm 1$, then H is free of rank d.*

Since in the case $\mu = 1$, H coincides with the commutator subgroup $[G, G]$ and the reduced Alexander polynomial $\tilde{\Delta}(t)$ becomes the ordinary Alexander polynomial $\Delta(t)$ of the knot, Theorem 1.1 implies immediately the following:

THEOREM 1.2. *Suppose K is an alternating knot. Then if $\Delta(0) = \pm 1$, $[G, G]$ of the knot group of K is free of rank d.*

Recently Neuwirth [6] and Rapaport [7] obtained some results on the commutator subgroup $[G, G]$ of the group G of a knot. Many of the arguments that were made in [6], [7] hold with slight modification in the case of the group of a link, if $[G, G]$ is replaced by H. Our proof of Theorem 1.1 based on these facts.

* Received November 13, 1962.

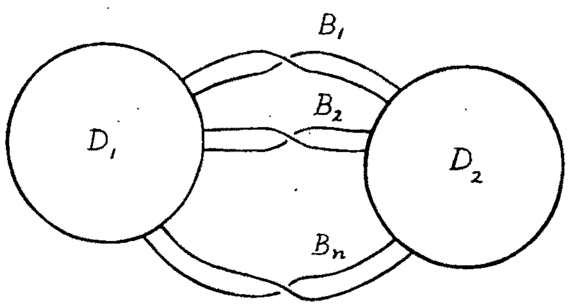
[1] Part of this paper was prepared while the author was at the University of Toronto, Canada.

Moreover, it should be noted that in Theorems 1.1 and 1.2, the hypothesis that L or K be alternating cannot be dropped [2].

In the following we do not distinguish exactly between knots and links. Therefore by a link (of multiplicity μ) is meant an ordinary knot or link according as $\mu=1$ or >1. Moreover, the reduced Alexander polynomial will be denoted by the same symbol $\Delta(t)$ as in the case $\mu=1$.

The author acknowledges his gratitude to Professor Fox for his helpful suggestions.

2. s-surfaces. Let us consider an orientable surface F in S^3, as is shown in Figure 1 below, consisting of two disks D_1, D_2 to which n bands $B_1, B_2, \cdots, B_n$ are attached. All B_i are twisted once in the same direction, and the bands are pairwise disjoint and do not link one another. Let us call F a *primitive s-surface* of type (n, ϵ), where $\epsilon = \pm 1$ according as the twisting is right-handed or left-handed. E.g. Figure 1 shows a primitive s-surface of type $(3, 1)$. The boundary $\dot{F}$ of F represents an elementary torus link of type $(2, n)$. In other words, F spans an elementary torus link.



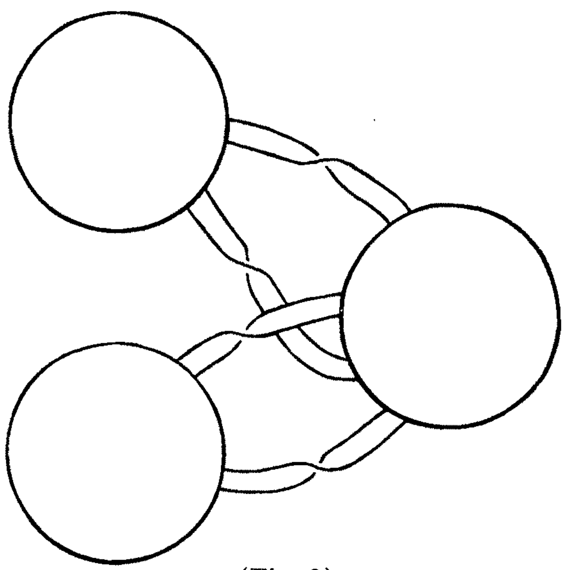
(Fig. 1)

Consider two primitive s-surfaces F and F' in S^3 of type (n, ϵ) and (m, η). Take two disks, D_1 and D_1' say, from each F and F' and identify them so that the resulting orientable surface $\tilde{F} = F \cup F'$ spans a link, and that $\tilde{F} - F$ and $\tilde{F} - F'$ are *separated*, i.e. there exists a 2-sphere S in S^3 such that $S \cap \tilde{F} = D_1$ $(= D_1')$ and each component of $S^3 - S$ contains points of $\tilde{F} - D_1$. $\tilde{F}$ consists of three disks and $n+m$ disjoint bands. $\tilde{F}$ will be called an *s-surface*. Similarly, given two s-surfaces, we can construct an orientable surface spanning a link by identifying two disks, one from each of the s-surfaces. In general, by an *s-surface* is meant an orientable surface obtained from a number of primitive s-surfaces by identifying their disks in this manner.

An identification of this kind leads to the product (in the sense of [4], [9]) of two given elementary torus links. Namely we have

(2.1) *Any product of two elementary torus links of type* $(2, m)$ *and* $(2, n)$ *is represented as the boundary of an s-surface with* $m+n$ *bands and three disks.*

As another example, Figure 2 shows an identification of two primitive s-surfaces of type $(2, 1)$ and $(2, -1)$ that produces a surface spanning the figure-eight knot.



(Fig. 2)

3. Alternating links.

LEMMA 3.1. *Any alternating link* L *for which* $\Delta(0) = \pm 1$, *can be spanned by an s-surface.*

Proof. Let p be a regular projection of L into S^2. Suppose that $p(L)$ is connected and *alternating* [1]. $p(L)$ possesses an orientation inherited from that of L. Now $p(L)$ is decomposed into a number of oriented circuits called *Seifert circuits,* in such a way that no two Seifert circuits have any side of $p(L)$ in common [1], [5]. A Seifert circuit C is said to be the *first* type if C bounds a 2-cell in $S^2 - p(L)$. Otherwise C is of the *second* type. Let us suppose that there are m Seifert circuits $C_1, \cdots, C_m$ of the second type. Since the underlying space $|C_i|$ of C_i is a simple closed curve in S^2, it divides S^2 into simply connected domains $|C_i|^\alpha$ and $|C_i|^\beta$. Let

$$R(\gamma_1, \cdots, \gamma_m) = \mathrm{Cl}(|C_1|^{\gamma_1} \cap \cdots \cap |C_m|^{\gamma_m}),$$

γ_i being α or β. Then only $m+1$ of the sets $R(\gamma_1, \cdots, \gamma_m)$ are non-empty. Let them be $R_1, \cdots, R_{m+1}$. R_i consists of some Seifert circuits of the second type, and two different domains R_i and R_j have at most one Seifert circuit of the second type in common.

Now $p^{-1}(p(L) \cap R_i)$ consists of some simple arcs in S^3. Their end points can be joined together in such a way that a special alternating link L_i results and $p(L_i) = p(L) \cap R_i$. We then write $L = L_1 * \cdots * L_{m+1}$. Since L is a link for which $\Delta(0) = \pm 1$, each L_i is a product of elementary torus links (Theorem 3.28 in [5]). Therefore, by (2.1), L_i is spanned by an s-surface F_i. Suppose R_i and R_j have a Seifert circuit C_k of the second type in common. Let r and s be the number of Seifert circuits contained in R_i and R_j. Then F_i and F_j contain just r and s disks, each of whose boundaries corresponds to a Seifert circuit. Let D and D' be disks of F_i and F_j corresponding to a common Seifert circuit C_k. Then we can identify D and D' in such a way that the resulting s-surface $F_i \cup F_j$ spans a link $L_i * L_j$. Thus, by performing the identification m times in this way, we obtain the required s-surface spannning the link $L = L_1 * \cdots * L_{m+1}$.

4. Proofs of theorems. Let S be an s-surface in S^3 and let U be an open regular neighborhood (in the sense of [11]) of the interior of S in S^3. U is the union of two surfaces $S^\#$ and S^b whose intersection is $\dot{S}$ and which span $\dot{S}$. Then the inclusion maps $\phi^\#: S^\# \to S^3 - U$ and $\phi^b: S^b \to S^3 - U$ induce homomorphisms

$$\phi^\#_*: \pi_1(S^\#) \to \pi_1(S^3 - U)$$

and

$$\phi^b_*: \pi_1(S^b) \to \pi_1(S^3 - U).$$

We claim, moreover, the following:

LEMMA 4.1. *$\phi^\#_*$ and ϕ^b_* are isomorphisms onto.*

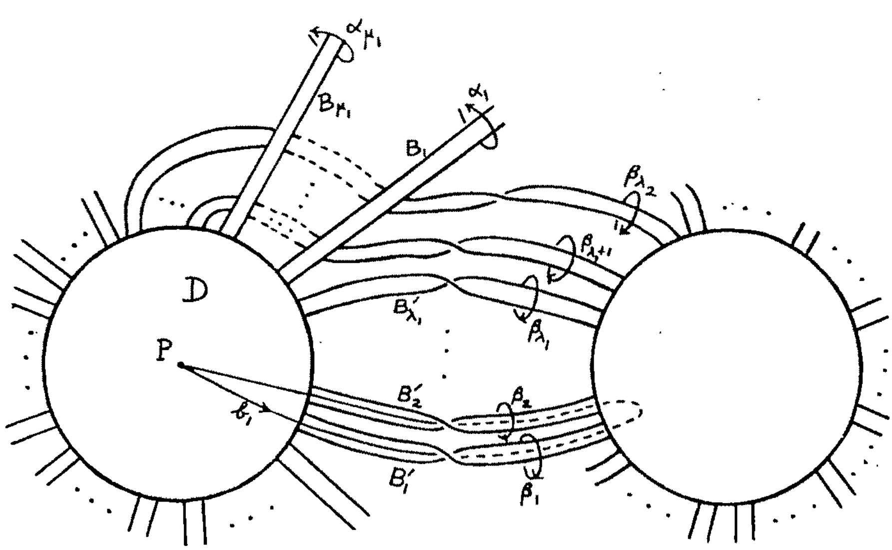
Proof. Proof need be made only for $\phi^\#_*$. If S is a primitive s-surface of type (n, ϵ), then $\phi^\#_*$ is clearly an isomorphism onto, and it is easy to show that $\pi_1(S^\#)$ and $\pi_1(S^3 - U)$ are free groups of the same rank $n-1$, and are freely generated by $b_1, \cdots, b_{n-1}$ and $\beta_1, \cdots, \beta_{n-1}$ respectively. Each b_i is represented by a simple closed curve on $S^\#$ that passes through two adjacent bands of $S^\#$, and each β_i is represented by a loop winding once around a band.

Now let S be an s-surface obtained from an s-surface F_1 and a primitive

s-surface F_2 of type (n, ϵ) by identifying their disks. For the sake of simplicity we suppose $\epsilon = 1$. Let U_i be open regular neighborhoods of the interiors of F_i in S^3. Because of the properties of the identification, we may take $U_1 \cup U_2$ as an open regular neighborhood U of S. Then $S^{\#} = F_1^{\#} \cup F_2^{\#}$, $S^b = F_1^b \cup F_2^b$, where $F_i^{\#}$, F_i^b have the meaning analogous to that of $S^{\#}$, S^b. We may assume inductively that

(4.1) (1) The inclusion map $\psi_i^{\#}: F_i^{\#} \to S^3 - U_i$ induces an isomorphism onto between their fundamental groups, $(i = 1, 2)$, and

(2) $\pi_1(F_1^{\#})$ and $\pi_1(S^3 - U_1)$ are free groups of rank r, generated by $a_1, \cdots, a_r$ and $\alpha_1, \cdots, \alpha_r$ respectively.



(Fig. 3)

Now $\pi_1(S^{\#})$ is a free product of two free groups $\pi_1(F_1^{\#})$ and $\pi_1(F_2^{\#})$ *amalgamated* on the subgroup $\pi_1(F_1^{\#} \cap F_2^{\#})$. Since this subgroup is trivial, $\pi_1(S^{\#})$ is the free group freely generated by $a_1, \cdots, a_r$ and $b_1, \cdots, b_{n-1}$. On the other hand, since $\pi_1(S^3 - U)$ is a *free product* of two free groups $\pi_1(S^3 - U_1)$ and $\pi_1(S^3 - U_2)$, $\pi_1(S^3 - U)$ is a free group freely generated by $\alpha_1, \cdots, \alpha_r$ and $\beta_1, \cdots, \beta_{n-1}$.

Let D be an identified disk of S and suppose that the bands attached to D are arranged in cyclic order as follows: $B'_1, \cdots, B'_{\lambda_1}\, B_1, \cdots, B_{\mu_1}, B'_{\lambda_1+1}, \cdots, B'_{\lambda_2}, \cdots, B'_{\lambda_{l-1}+1}, \cdots, B'_{\lambda_l}, B_{\mu_{l-1}+1}, \cdots, B_{\mu_l}$, where B_i and B'_j are bands of F_1 and F_2.

Let α_i $(i=1,\cdots,\mu_l)$ and β_j $(j=1,2,\cdots,\lambda_l=n-1)$ be elements represented by loops winding once around B_i and B_j'. We may assume, moreover, without loss of generality, that there is a regular projection of L whose image in the neighborhood of D will be shown in Figure 3. (We have only to deform L isotopically, if necessary.)
Then by choosing the base point P of $\pi_1(S^{\#})$ in $F_1^{\#}\cap F_2^{\#}$, the induced homomorphism $\phi^{\#}_*$ is given by

$$(4.2)\qquad \phi^{\#}_*(a_i)=\psi^{\#}_{1*}(a_i),\ i=1,2,\cdots,r,$$
$$\phi^{\#}_*(b_j)=\begin{cases}\beta_j, & \text{for } 1\leqq j\leqq\lambda_1-1,\\ (A_{k-1}^{-1}A_{k-2}^{-1}\cdots A_1^{-1})\beta_j(A_1\cdots A_k), & \text{for } j=\lambda_k,\ k=1,\cdots,l-1,\\ (A_k^{-1}\cdots A_1^{-1})\beta_j(A_1\cdots A_k), & \text{for } \lambda_k+1\leqq j\leqq\lambda_{k+1}-1,\\ & k=1,\cdots,l-1.\end{cases}$$

where $A_h=\alpha_{\mu_{h-1}+1}\alpha_{\mu_{h-1}+2}\cdots\alpha_{\mu_h}$, for $h=1,\cdots,l-1$, $\mu_0=0$, and $A_0=1$. Thus since by the assumption of induction $\psi^{\#}_{1*}$ is an isomorphism onto, it follows that $\phi^{\#}_*$ is also an isomorphism onto.

As remarked in § 1, Lemma 4.1, and a slight modification of Theorem 1 in [6] imply that

LEMMA 4.2. *H is a free group.*

Moreover, if H is free, then a slight modification of results of Theorem 1 and Theorem 3, Corollary in [7] also show that the rank of H must be the degree of $\Delta(t)$ and $\Delta(0)=\pm1$. Thus we obtain the following

THEOREM 4.3. *Let S be an s-surface and let d be the degree of the reduced Alexander polynomial $\Delta(t)$ of a link $\dot{S}$. Then H is free of rank d and $\Delta(0)=\pm1$.*

Theorems 1.1 and 1.2 follow from Lemma 3.1 and Theorem 4.3.

Remark 1. Any s-surface S is one of the surfaces with minimum genus spanning the oriented link $\dot{S}$. Because, let h be the genus of S. Since H is free and inclusion maps $S^{\#}\to(S^3-U)$ and $S^b\to(S^3-U)$ induce isomorphisms onto, the proof of Theorem 1 in [6] can be modified to hold in the case of the group of a link. Thus the rank d of H must be $2h+\mu-1$. On the other hand, since $d\leqq 2g+\mu-1$, g denoting the *genus* of $\dot{S}$, we see $d\leqq 2g+\mu-1\leqq 2h+\mu-1=d$. Therefore $g=h$. d is equal to $N-M+1$, where N and M denote the number of bands and disks of S.

Remark 2. Generally the converse of Theorem 4.3 does not hold. For example, 9_{44} in [8] cannot be spanned by any s-surface, but its group G and polynomial $\Delta(t)$ satisfy the conclusion in Theorem 4.3.

HOSEI UNIVERSITY, JAPAN,
AND PRINCETON UNIVERSITY.

REFERENCES.

[1] R. H. Crowell, "Genus of alternating link types," *Annals of Mathematics*, vol. 69 (1959), pp. 259-275.

[2] ——— and H. F. Trotter, "A class of Pretzel knot," to appear in *Duke Mathematical Journal*.

[3] R. H. Fox, "Free differential calculus, II," *Annals of Mathematics*, vol. 57 (1953), pp. 547-560.

[4] Y. Hashizume, "On the uniqueness of the decomposition of a link," *Osaka Mathematical Journal*, vol. 10 (1958), pp. 283-300.

[5] K. Murasugi, "On alternating knots," *Osaka Mathematical Journal*, vol. 12 (1960), pp. 277-303.

[6] L. Neuwirth, "The algebraic determination of the genus of knots," *American Journal of Mathematics*, vol. 82 (1960), pp. 791-798.

[7] E. Rapaport, "On the commutator subgroup of a knot group," *Annals of Mathematics*, vol. 71 (1960), pp. 157-162.

[8] K. Reidemeister, *Knotentheorie*, Berlin, 1932.

[9] H. Schubert, "Die eindeutige Zerlegbarkeit eines Knotens in Primknoten," *Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Mathematische-naturwissenschaftliche Klasse*, 1949, pp. 57-104.

[10] H. Seifert, "Ueber das Geschlecht von Knoten," *Mathematische Annalen*, vol. 110 (1934), pp. 571-592.

[11] J. H. C. Whitehead, "Simplicial spaces, nuclei and m-groups," *Proceedings of the London Mathematical Society*, vol. 45 (1939), pp. 243-327.

ON GROUPS OF MEASURE PRESERVING TRANSFORMATIONS. II.*

By H. A. Dye.[1]

1. Introduction. In this paper we resume the study, begun in [3], of automorphism groups of finite non-atomic measure alegbras, the main object of the theory being the classification of such groups up to weak equivalence. Each group G of measure preserving automorphisms is immersed in a full group $[G]$, consisting of all automorphisms α admitting a representation $\alpha(E) = \sum_{g \in G} Q_g g(E)$, for appropriate elements Q_g of the measure algebra, and two groups are called weakly equivalent if their corresponding full groups are conjugate, in the usual sense of ergodic theory. A striking feature of this equivalence concept is its coarseness. It is a common occurrence for groups with no familiar algebraic features in common to be weakly equivalent, and similarly, diverse action characteristics do not affect weak equivalence. Perhaps significantly, we as yet possess no examples of an abstract group having two inequivalent faithful repersentations as freely acting ergodic automorphism groups of a separable non-atomic measure algebra.

The main results in [3] concerned approximately finite groups: an automorphism group G is called approximately finite if elements in each finite subset of G can be approximated arbitrarily closely in the natural metric of $[G]$ by elements in an appropriate finite subgroup of $[G]$. In this form, the definition of approximate finiteness leads rather quickly to a structure and equivalence theory. However, it is usually a major technical problem to verify whether a given automorphism group satisfies this definition. This problem serves as the point of departure for the present paper. Specifically, we consider a freely acting automorphism group G which is generated by a symmteric finite subset F, and attempt to cast the definition of approximate finiteness in an equivalent form more accessible in testing examples. An equivalent form is found, in Proposition 4.1, with which we are able to establish a criterion (Theorem 1) for approximate finiteness of G in terms of its group structure alone. This is a statement that if the powers F^n of F do not grow too rapidly, in a certain sense, then G is approximately finite.

* Received August 27, 1962.

[1] This research was supported in part by NSF grant 18999.

The criterion is satisfied by all abelian groups, and as we point out in Section 6, this leads to a generalization of a result of Murray-von Neumann on approximately finite factors.

In Section 5, carrying on the same theme—characterization of the weak equivalence type of an automorphism group by its abstract group structure alone—, we show that two full groups of type II are weakly equivalent if and only if they are isomorphic as groups (Theorem 2). Our technique in proof is to develop a structure theory for such groups sufficient to show that local subgroups G_P of a full group G of type II can be distinguished internally in G as an abstract group.

In an appendix to this paper, Section 6, we discuss certain full groups of measure preserving automorphisms which arise in the study of regular maximal abelian subalgebras of finite von Neumann algebras. The approximate finiteness of these groups entails the approximate finiteness of the corresponding von Neumann algebras, so that a basis is provided for application of certain results of this theory to the study of von Neumann algebras.

2. Technical preliminaries. We give a brief resume of the terminology developed in [3]. The objects under study are groups of measure preserving ($= MP$) automorphisms of finite non-atomic measure algebras. For technical reasons, it is convenient to deal with what we term "hyperstonian measure spaces" raher than with general measure algebras (though nothing is lost, for with each finite measure algebra is associated a uniquely determined hyperstonian measure space having the same measure algebra. A hyperstonian measure space (M, λ) is a pair consisting of (1) an abelian C^*-algebra M with identity I, in which each bounded collection of self-adjoint elements has a least upper bound in M, relative to the usual order, and (2) a positive linear function λ on M with $\lambda(I) = 1$ which is faithful, in the sense that the only positive element in the kernel of λ is 0, and normal, in the sense that $\lambda(\mathrm{LUB}_\alpha A_\alpha) = \mathrm{LUB}_\alpha \lambda(A_\alpha)$, for each bounded collection of self-adjoint elements A_α closed under the formation of finite joins. M_P denotes the collection of projections ($=$ self-adjoint idempotents) in M; we assume throughout that the measure algebra (M_P, λ) is non-atomic.

Given a group G of MP automorphisms of (M, λ), we denote by Z_G the fixed algebra of G, that is, the collection of all elements in M fixed under each automorphism in G. A projection P is called abelian over Z_G if $P \neq 0$ and each projection $Q \leqq P$ has the form $Q = PC$, for some C in Z_G. G is of type I if each non-zero projection dominates a projection abelian over Z_G, and G is of type II if no projections abelian over Z_G exist. If α and β are

MP automorphisms, then $F(\alpha, \beta)$ denotes the maximal projection F such that $\alpha(P) = \beta(P)$, for all $P \leqq F$. The full group $[G]$ determined by G is the collection of all MP automorphisms α such that $\mathrm{LUB}_{g \in G} F(\alpha, g) = I$, and a group G is called full if $[G] = G$. On the other hand, if $F(g, h) = 0$ for each pair $g \neq h$ in G, then G is called freely-acting. If K is a full subgroup of $[G]$ and $g \in G$, then $E(K, g)$ denotes the projection $\mathrm{Lub}_{k \in K}\, gF(g, k)$. The MP automorphism group G is called approximately finite if, for each finite subset F of G and each $\epsilon > 0$, there exists a type I subgroup K of $[G]$ such that $\lambda(E([K], x)) > 1 - \epsilon$, for all $x \in F$. (An equivalent definition results here if one replaces "type I subgroup" by "finite subgroup.")

Given an MP group G on (M, λ) and an MP group G' on (M', λ'), we say that G is weakly equivalent to G' if there exists an isomorphism θ of M on M' such that $[\theta G \theta^{-1}] = [G']$. Here isomorphism means only C^*-algebra isomorphism; θ is not required to be measure preserving. In any case, however, one will have $\theta(E_{Z_G}(A)) = E_{Z_{G'}}(\theta(A))$, for all $A \in M$, where $E_{Z_G}(A)$ denotes the conditional expectation of A relative to the fixed algebra Z_G of G, and so, for example, θ will conserve measure when the groups are ergodic.

3. Determining functions. Let G be a group of MP automorphisms of the non-atomic hyperstonian measure space (M, λ). By a *determining function* on G we mean a mapping $x \to E_x$ of G into M_P satisfying

$$E_x E_{xy} = E_x x E_y \text{ (for all } x, y) \text{ and } E_e = I. \tag{3.1}$$

For example, if P is a fixed projection, and if one sets $E_x = PxP$ $(x \neq e)$ and $E_e = I$, then E_x satisfies (3.1). Similarly, $E_x = F(x, e)$ is a determining function. There are several useful elementary observations: (1) any determining function satisfies $xE_{x^{-1}} = E_x$ (as one sees by substituting $y = x^{-1}$ in (3.1) and using a symmetry argument): (2) if E_x and F_x are determining functions, then $E_x F_x$ is a determining function; (3) determining functions on a subgroup of a group can be extended to the group in various ways. There is the trivial extension: if E_x is defined on the subgroup H of G, and if one sets $F_x = E_x$ (x in H) and $F_x = 0$ (x in G not H), then F_x is a determining function on G. On the other hand, suppose G is freely-acting. Then each α in $[G]$ has a unique representation $\alpha = \sum_{x \in G} Q(\alpha, x) x$ (where $Q(\alpha, x) = xF(\alpha, x)$). Now if E_x is a determining function on G, then the function $E_\alpha = \sum_{x \in G} Q(\alpha, x) E_x$ is a determining function on $[G]$, which we shall term the natural extension of E_x to $[G]$.

Given a determining function E_x on G, we associate with E_x a full sub-

group $K(E_x)$ of $[G]$ in the following way. Let $K=K(E_x)$ be the set of all α in $[G]$ which admits a representation $\alpha=\sum_{x\in G} Q_x x$ with $Q_x\leqq E_x$ for all x. K is a subgroup of $[G]$: if $\alpha=\sum Q_x x$ and $\beta=\sum R_x x$ with $Q_x\leqq E_x\geqq R_x$, then $\alpha^{-1}\beta=\sum_y[\sum_x x^{-1}Q_x x^{-1}R_{xy}]y$, and

$$\sum_x x^{-1}Q_x x^{-1}R_{xy}=\sum_x x^{-1}Q_x x^{-1}R_{xy}x^{-1}E_x x^{-1}E_{xy}=\sum_x x^{-1}Q_x x^{-1}R_{xy}E_{x^{-1}}E_y\leqq E_y,$$

showing that $\alpha^{-1}\beta\in K$. And K is full: each α in $[K]$ has a representation $\alpha=\sum_n R_n\alpha_n$ with $\alpha_n=\sum_{x\in G} Q(n,x)x$, $Q(n,x)\leqq E_x$, so that

$$\alpha=\sum_x[\sum_n R_nQ(n,x)]x \text{ and } \sum_n R_nQ(n,x)\leqq\sum_n R_nE_x=E_x.$$

We will call $K(E_x)$ the subgroup of $[G]$ determined by E_x.

Conversely, let K be a full subgroup of $[G]$. Then the function $E(K,x)=\mathrm{LUB}_{k\in K}\, xF(x,k)$ is a determining function:

$$\begin{aligned}E(K,x)xE(K,y)&=\mathrm{LUB}_{h,k\in K}xF(x,h)xyF(y,k)\\&=\mathrm{LUB}_{h,k\in K}xF(x,h)xyF(y,h^{-1}k)=\mathrm{LUB}_{h,k\in K}xF(x,h)xyF(hy,k)\\&=\mathrm{LUB}_{h,k\in K}xF(x,h)xyF(xy,k)=E(K,x)E(K,xy),\end{aligned}$$

and clearly, $E(K,e)=I$.

LEMMA 3.1. *For any full subgroup K of $[G]$, $K(E(K,x))=K$. For any determining function E_x on G, $E_x\leqq E(K(E_x),x)$, equality holding for all x if and only if $E_x\geqq F(x,e)$ for all x.*

Proof. If $\alpha=\sum_{x\in G} Q_x x$, with $Q_x\leqq E(K,x)$ for all x, then α lies in K: for, by [3, Lemma 3.4], there exist elements β_x in K such that $Q_x\beta_x=Q_x x$, and because K is full, $\alpha=\sum Q_x\beta_x$ must lie in K. On the other hand, if $\alpha=\sum_{x\in G} Q_x x$ is *any* representation of $\alpha\in K$ in terms of G, then $Q_x\alpha=Q_x x$ for all x, and the lemma just cited forces $Q_x\leqq E(K,x)$. Therefore, $K(E(K,x))=K$.

Given a determining function E_x, write $K(E_x)=K$. We can express $(I-F(x,e))E_x$ in the form $\sum_n P_n$, where the P_n are mutually orthogonal and $P_nx^{-1}P_n=0$ for each n. The mapping $P_nx+x^{-1}P_nx^{-1}+(I-(P_n+x^{-1}P_n))e$ is then an element of $K(E_x)$, and by the preceding paragraph, $P_n\leqq E(K,x)$. Therefore, $(I-F(x,e))E_x\leqq E(K,x)$. Already $F(x,e)\leqq E(K,x)$, so it follows that $E_x\leqq E(K,x)$. Now assume that $E_x\geqq F(x,e)$ for all x. We claim that $E(K,x)\leqq E_x$. In fact, take any P such that $Px=P\alpha$, for fixed x in G and α in K. If $\alpha=\sum_{y\in G} Q_yy$, then $Q_yPx=Q_yPy$ for each y. This entails

$Q_yP \leqq yF(x,y) = yF(y^{-1}x, e)$, and since $F(y^{-1}x, e) \leqq E_{y^{-1}x}$ and $Q_y \leqq E_y$, we have $Q_yP \leqq E_y y E_{y^{-1}x} = E_y E_x \leqq E_x$. Consequently, $P = \sum Q_y P \leqq E_x$, and we have $E(K,x) \leqq E_x$, completing the proof.

We proceed now with the assumption that G is freely-acting. If E_x is any determining function on G, then clearly $E_x \geqq F(x,e)$ for all x, so by the lemma, $E_x \to K(E_x) \to E(K(E_x),x)$ establishes a mutual $1:1$ correspondence between determining functions on G and full subgroups of $[G]$. It is not hard to identify the subgroup corresponding to the product of two determining functions E_x and F_x, nor the subgroup corresponding to PxP (P fixed). In fact,

$$K(E_xF_x) = K(E_x) \cap K(F_x), \tag{3.2}$$

and

$$PxP = E([G]_P, x) \quad (x \neq e). \tag{3.3}$$

(3.2) is a restatement of [3, Lemma 5.5]. To prove (3.3), take α in $[G]_P$ $(= [\alpha \in [G] \mid F(\alpha, e) \geqq I - P])$, $\alpha = \sum_{x \in G} Q(\alpha, x)x$. If $x \neq e$, the free action of G forces $F(x,\alpha) \leqq P$, and we have $Q(\alpha,x) = xF(x,\alpha) = \alpha F(x,\alpha)$. So on the one hand, $Q(\alpha,x) = xF(x,\alpha) \leqq xP$, and on the other $Q(\alpha,x) = \alpha F(x,\alpha) \leqq \alpha P = P$, forcing $Q(\alpha,x) \leqq PxP$. Conversely, if α lies in the subgroup of $[G]$ determined by PxP, we can write $\alpha = \sum Q_x x$ with $Q_x \leqq PxP$ for $x \neq e$. Then if $Q \leqq I - P$, $\alpha(Q) = Q_eQ + \sum_{x \neq e} Q_x PxPxQ = Q_eQ$, forcing $\alpha(Q) = Q$. This implies that $\alpha \in [G]_P$.

Let E_x be a determining function on the freely-acting group G, and write $K = K(E_x)$. We associate with E_x the extended real-valued function $E = \sum_x E_x$ $(= \mathrm{LUB}_F \sum_{x \in F} E_x$, F ranging over finite subsets of G) on the spectrum of M, which we call the *type function* of K. When E is bounded, we shall consider it as an element of M. We call E finite almost everywhere if there exists an increasing sequence Q_n of projections in M with LUB I such that each $\sum_x E_x Q_n$ is bounded. (An equivalent definition is this: E is finite almost everywhere if each projection $Q \neq 0$ dominates a $P \neq 0$ such that $PE_x = 0$ for all except finitely many x.)

PROPOSITION 3.1. *In order that K be of type I, it is necessary and sufficient that its type function be finite almost everywhere. When the type function E is finite a. e., there exist mutually orthogonal projections D_n in the fixed algebra Z_K of K such that $\sum_n D_n = I$, K_{D_n} is of type I_n on D_nM, and $E = \sum nD_n$.*

Proof. Observe first that $Z_K = [A \in M \mid E_x A = E_x x A$ for all x in $G]$. In fact, if $E_x A = E_x x A$ for all x and if $\alpha = \sum_x Q(\alpha, x) x$ lies in K, then we have $Q(\alpha, x) \leqq E_x$, so that $\alpha(A) = \sum Q(\alpha, x) E_x x A = \sum Q(\alpha, x) E_x A = A$, and $A \in Z_K$. On the other hand, since $E(K, x) = \mathrm{LUB}_{\alpha \text{ in } K} Q(\alpha, x)$, A in Z_K entails $Q(\alpha, x) x A = Q(\alpha, x) A$ for all α in K, and $E_x x A = E_x A$ follows because $E_x = E(K, x)$.

Assume now that $E = \sum E_x$ is bounded, and write $\phi(A) = E^{-1} \sum E_x x A$. Then

$$\phi(A) = E_{Z_K}(A), \tag{3.4}$$

where E_{Z_K} denotes conditional expectation relative to Z_K. For this we have to show that $\phi(A) \in Z_K$, for all A in M, and that $\lambda(\phi(A)B) = \lambda(AB)$, for all B in Z_K. First, note that $E \in Z_K : E_y y E = \sum_x E_y y E_x = \sum_x E_y E_{yx} = \sum_x E_y E_x = E_y E$, and by the preceding paragraph, $E \in Z_K$. In the same way, it follows that $E_y y \phi(A) = E_y \phi(A)$, for any A in M, whence $\phi(A) \in Z_K$. Finally, $\lambda(\phi(A)B) = \sum_y \lambda(E^{-1} E_y (yA) B) = \sum_y \lambda((y^{-1} E_y) E^{-1} AB) = \sum_y \lambda(E_{y^{-1}} E^{-1} AB) = \lambda(AB)$, proving (3.4).

By the structure theory in [3, Section 4], there exist uniquely determined mutually orthogonal projections $D_\infty, D_1, D_2, \cdots$ in Z_K having LUB I, and such that K_{D_n} is of type I_n on $D_n M$ (n finite) and K_{D_∞} is of type II on $D_\infty M$.

Assume that $D_\infty \neq 0$. We will show that for no $0 \neq P \leqq D_\infty$ is $\sum P E_x$ bounded. Using an indirect proof, assume to the contrary that $\sum P E_x \leqq mI$ for some integer m and some $0 \neq P \leqq D_\infty$. Put $F_x = P x P E_x$ $(x \neq e)$, $F_e = I$. Then $\sum F_x \leqq mI$ and, by (3.2) and (3.3), $F_x = E(K_P, x)$. Let L denote the fixed algebra of K_P. Then by (3.4), $(\sum F_x)^{-1} \sum F_x x A = E_L(A)$. Now K_P is of type II on PM, and so by Maharam's lemma [3, p. 124], there exists a projection $Q \leqq P$ such that $E_L(Q) = P/2m$. It follows that

$$\sum_x F_x x Q = (\sum F_x) P/2m \leqq P/2.$$

By the nature of the left-hand term, this forces $\sum F_x x Q = 0$, $E_L(Q) = 0$, a contradiction. Therefore, if E is finite a. e., then K has no type II summands.

Now assume that K is of type I (that is, $D_\infty = 0$). There exists a freely-acting group K_n of order n acting on $D_n M$ which is equivalent to K_{D_n} on $D_n M$. Each k in K_n has a unique representation $k = \sum_{x \in G} Q(k, x) x$ with $\sum_x Q(k, x) = D_n$. Now $E(K, x) D_n = E([K_n], x) D_n = \sum_{k \in K_n} Q(k, x)$, so that

$$\sum_x E_x D_n = \sum_x [\sum_{k \in K_n} Q(k, x)] = \sum_{k \in K_n} [\sum_x Q(k, x)] = \sum_{k \in K_n} D_n = n D_n.$$

It follows that $E = \sum_x E_x = \sum_n nD_n$ is finite a.e., and the lemma is proved.

In order, therefore, that the subgroup K of the proposition be of type II, it is necessary and sufficient that, for each $P \neq 0$, one have $PE_x \neq 0$ for infinitely many x.

Continuing in the case G freely-acting, we give now an explicit method for construction of determining functions which is useful in the study of approximately finite groups. Let F be a finite subset of G which is *symmetric* in the sense that $e \in F = F^{-1}$. The subgroup H of G generated by F is then simply $\bigcup_{n=1}^{\infty} F^n$. We suppose that a mapping $x \to P(x)$ of F into M_P is given which satisfies

$$(3.5) \qquad P(e) = I, \qquad P(x) = xP(x^{-1}).$$

(For example, if F_x is a determining function on G, then the mapping $P(x) = F_x$, defined for x in F, satisfies (3.5).) For each integer $k \geqq 1$, write $F_k = F$ and form the cartesian products $F^{(m)} = \prod_{k=1}^{m} F_k$. Denote by R_x the set of all $(x_1, \cdots, x_n)$ in $\bigcup_{m=1}^{\infty} F^{(m)}$ such that $x_1 \cdots x_n = x$. In turn, denote by R_x^0 those $(x_1, \cdots, x_n)$ in R_x with the additional property that $x_1, x_1x_2, \cdots, x_1 \cdots x_n$ $(= x)$ are distinct elements of G. For $(x_1, \cdots, x_n)$ in R_x, define

$$(3.6) \qquad P(x_1, \cdots, x_n) = P(x_1)x_1P(x_2) \cdots (x_1x_2 \cdots x_{n-1})P(x_n),$$

and finally, define

$$(3.7) \qquad E_x = \mathrm{LUB}_{(x_1, \cdots, x_n) \text{ in } R_x} P(x_1, \cdots, x_n).$$

LEMMA 3.2. (i) *E_x is a determining function on H*; (ii)

$$E_x = \mathrm{LUB}_{(x_1, \cdots, x_n) \in R_x^0} P(x_1, \cdots, x_n);$$

(iii) *if the original function $P(x)$ arises by restriction of a determining function F_x to F, then $E_x \leqq F_x$, equality holding on F*; (iv) *in order that the type function E of E_x be bounded, it is necessary and sufficient that E_x vanish off a finite subset of H.*

Proof. (i) From (3.5) and (3.6), one has the identities

$$P(x_n^{-1}, \cdots, x_1^{-1}) = (x_n^{-1} \cdots x_1^{-1})P(x_1, \cdots, x_n)$$

and

$$P(x_1, \cdots, x_n, x_{n+1}, \cdots, x_{n+m}) \\ = P(x_1, \cdots, x_n)(x_1x_2 \cdots x_n)P(x_{n+1}, \cdots, _{n+m}).$$

So, if $(x_1, \cdots, x_n) \in R_x$ and $(z_1, \cdots, z_m) \in R_{xy}$, then
$P(x_1, \cdots, x_n)P(z_1, \cdots, z_m)$

$$= P(x_1, \cdots, x_n)x_1x_2 \cdots x_n [P(x_n^{-1}, \cdots, x_1^{-1})x_n^{-1} \cdots x_1^{-1}P(z_1, \cdots, z_m)]$$
$$= P(x_1, \cdots, x_n)x_1x_2 \cdots x_n P(x_n^{-1}, \cdots, x_1^{-1}, z_1, \cdots, z_m) \leqq E_x x E_{x^{-1}xy}$$
$$= E_x x E_y.$$

It follows that $E_x E_{xy} \leqq E_x x E_y$. On the other hand, if $(x_1, \cdots, x_n) \in R_x$ and $(y_1, \cdots, y_m) \in R_y$, then

$$P(x_1, \cdots, x_n)xP(y_1, \cdots, y_m)$$
$$= P(x_1, \cdots, x_n)P(x_1, \cdots x_n, y_1, \cdots, y_m) \leqq E_x E_{xy}.$$

Therefore, $E_x x E_y \leqq E_x E_{xy}$. Obviously $E_e = 1$, and it follows that E_x satisfies (3.1) and is therefore a determining function on H.

(ii) Suppose $(x_1, \cdots, x_n) \in R_x$ and that the repetition

$$x_1 \cdots x_r = x_1 \cdots x_r \cdots x_{r+s} \quad (r, s \geqq 1)$$

occurs in $x_1, x_1x_2, \cdots, x_1x_2 \cdots x_n$. Then $x_{r+1} \cdots x_{r+s} = e$ and

$$(x_1, \cdots, x_r, x_{r+s+1}, \cdots, x_n) \in R_x.$$

We claim that $P(x_1, \cdots, x_n) \leqq P(x_1, \cdots, x_r, x_{r+s+1}, \cdots, x_n)$. In fact,
$P(x_1, \cdots, x_n)$

$$= P(x_1, \cdots, x_r)(x_1x_2 \cdots x_r) [P(x_{r+1}, \cdots, x_{r+s})x_{r+1} \cdots x_{r+s}P(x_{r+s+1}, \cdots, x_n)]$$
$$= P(x_1, \cdots, x_r)x_1 \cdots x_r[P(x_{r+1}, \cdots, x_{r+s})P(x_{r+s+1}, \cdots, x_n)]$$
$$\leqq P(x_1, \cdots, x_r)x_1 \cdots x_r P(x_{r+s+1}, \cdots, x_n)$$
$$= P(x_1, \cdots, x_r, x_{r+s+1}, \cdots, x_n).$$

Continuation of this process for a finite number of steps will eliminate repetitions and give the inequality $E_x \leqq \mathrm{LUB}_{(x_1, \cdots, x_n) \in R_x^0} P(x_1, \cdots, x_n)$. But the inequality $\geqq$ is automatic, and (ii) follows.

(iii) Given $F_x = P(x)$ (x in F), we have

$$P(x_1, \cdots, x_n) = F_{x_1}x_1F_{x_2} \cdots (x_1 \cdots x_{n-1}) F_{x_n} = F_{x_1}F_{x_1x_2} \cdots F_{x_1 \cdots x_n},$$

making repeated use of (3.1). This shows that $P(x_1, \cdots, x_n) \leqq F_{x_1 \cdots x_n}$. Therefore, putting $x = x_1 \cdots x_n$, $E_x \leqq F_x$. If $x \in F$, then $F_x = P(x) \leqq E_x$, so $F_x = E_x$.

(iv) Suppose that the type function is bounded by the integer n. Let x be any element of H not in $\bigcup_{k=1}^{n} F^k$, and suppose $E_x \neq 0$. Then $P(x_1, \cdots, x_m)$

$\neq 0$, for some $(x_1, \cdots, x_m)$ in R_x^0, and by the nature of x, $m > n$. But $P(x_1, \cdots, x_m) = P(x_1, x_2) \cdots P(x_1, \cdots, x_m) \leqq E_{x_1} E_{x_1 x_2} \cdots E_{x \cdots x_m}$. This shows that $\sum_x E_x \geqq m$ on a non-void clopen set, contradicting $\sum_x E_x \leqq n < m$. It follows that E_x vanishes off the finite set $\bigcup_{k=1}^{n} F^k$. This proves the lemma.

4. Approximate finiteness. Using the methods of Section 3, we now undertake to determine conditions on a freely-acting group G, both in terms of its action and its algebraic structure, which will entail the approximate finiteness of G.

Let F be a symmetric finite subset of G. We say that a projection P is *F-dispersed* if there exists an integer n_0 such that, for each $n > n_0$ and each n-tuple $(x_1, \cdots, x_n)$ $(x_i \in F)$ with $x_1, x_1 x_2, \cdots, x_1 \cdots x_n$ distinct (viz. $(x_1, \cdots, x_n) \in R_{x_1 \cdots x_n}{}^0$), one has $P \cup x_1 P \cup \cdots \cup x_1 x_2 \cdots x_n P = I$.

PROPOSITION 4.1. *In order that the freely-acting group G be approximately finite, it is necessary and sufficient that, for each symmetric finite subset F of G, there exist F-dispersed projections of arbitrarily small measures.*

Proof. Assume G approximately finite, and let F be a symmetric finite subset. Given $\epsilon > 0$, we can choose a full bounded type I subgroup K of $[G]$ such that $\lambda(P) > 1 - \epsilon$, where we set $\prod_{x \in F} E(K, x) = P$. We contend that $I - P$ is F-dispersed. To see this, put $P(x) = E(K, x)$ (x in F), and let F_x be the determining function constructed from the function $P(x)$ on F by the procedure of Lemma 3.2. By that lemma we have $F_x = E(K, x)$ for x in F, and $F_y \leqq E(K, y)$ for all y. It follows that $P = \prod_{x \in F} F_x$ and that the type function of F_x is bounded. Therefore, by Lemma 3.2, F_x has finite support. It follows that there exists an n_0 such that $F_{x_1 \cdots x_n} = 0$ whenever $(x_1, \cdots, x_n) \in R_{x_1 \cdots x_n}{}^0$ and $n > n_0$. For each n-tuple,

$$P x_1 P \cdots x_1 \cdots x_n P = P(x_1, \cdots, x_n) \leqq F_{x_1 \cdots x_n} = 0.$$

Taking complements, it follows that $I - P$ is F-dispersed.

Assume conversely that for each symmetric finite subset F of G, there exist F-dispersed projections of arbitrarily small measure. Given a (symmetric) finite subset F and an $\epsilon > 0$, we want to construct a type I subgroup K of $[G]$ such that $\lambda(E([K], x)) > 1 - \epsilon$, for all x in F.

For this, let $I - P$ be an F-dispersed projections with $\lambda(I - P) < \epsilon/2$. Set $P(x) = PxP$ (x in F, $x \neq e$), $P(e) = I$, and let E_x be the determining function constructed from this $P(x)$ by Lemma 3.2. By the choice of P,

there exists an n_0 such that $P(x_1,\cdots,x_n)=0$ whenever $(x_1,\cdots,x_n)$ lies in $R_{x_1\cdots x_n}{}^0$ and $n>n_0$. It follows that $E_{x_1\cdots x_n}=0$ for such n-tuples. We conclude that E_s has finite support, and therefore it determines a type I subgroup of $[G]$. Since $E_s\geqq PxP$ for x in F, we have $\lambda(E_s)>1-\epsilon$, and the proposition is proved.

From the definition of approximate finiteness, it is clear that a group is approximately finite if each finitely generated subgroup is approximately finite. In turn, if the group G is generated by a finite subset F, and if the condition defining approximate finiteness holds for elements in this particular finite subset F, then it is straightforward to show that G is approximately finite. Therefore, if the freely-acting group G is generated by a symmetric finite subset F, and if there exist F-dispersed projections of arbitrarily small measure, then G is approximately finite.

Given a subset S of G and a projection P, write SP for $LUB_{s\in S}\ sP$.

LEMMA 4.1. *Let S be an arbitrary finite subset of the MP group G, and let Q be an arbitrary projection. Let*

$$C=\bigcap_{x\in F}xS \text{ and } P=(CQ)(I-(S-C)Q),$$

F being a symmetric finite subset of G. Then $I-P$ is F-dispersed.

Proof. Suppose that $x_1,x_1x_2,\cdots,x_1\cdots x_n$ $(x_i\in F)$ are distinct and $Px_1P_1\cdots x_1\cdots x_nP\neq 0$. We will show that n cannot exceed the number $|C|$ of elements in C. In fact, the projection $Px_1P\cdots x_1\cdots x_nP$ must have a non-zero intersection R with some c_0Q (c_0 in C). We claim that $x_1^{-1}c_0\in C$; for by definition of C, $x_1^{-1}c_0\in S$; but $x_1^{-1}R\leqq P$, which is disjoint from $(S-C)Q$, and because $x_1^{-1}R\leqq x_1^{-1}c_0Q\leqq SQ$, we must have $x_1^{-1}c_0\in C$. Similarly, $x_2^{-1}x_1^{-1}c_0\in S$ and $x_2^{-1}x_1^{-1}R\leqq Px_2^{-1}x_1^{-1}c_0Q$, forcing $x_2^{-1}x_1^{-1}c_0\in C$. It follows that $x_i^{-1}\cdots x_1^{-1}c_0\in C$ for $i=1,\cdots,n$. Since these elements are all distinct, we conclude that $n\leqq|C|$. Therefore, if $n>|C|$ and

$(x_1,\cdots,x_n)\in R_{x_1\cdots x_n}{}^0$, then $Px_1P\cdots x_1\cdots x_nP=0$. It follows that $I-P$ is F-dispersed.

THEOREM 1. *Let G be a freely-acting group generated by a symmetric finite subset F. Define $h_1=|F|$, $h_n=|F^n-F^{n-1}|$ $(n>1)$. Then, if*

$$\inf_n h_{2n}/(h_1+\cdots+h_n)=0. \tag{4.1}$$

the group G is approximately finite.

Proof. Given $\epsilon>0$, choose n so that $h_{2n}<\epsilon(h_1+\cdots+h_n)$. Put

$A = F^n$ and define $S = F^{2n}$ $(= A^2 = A^{-1}A)$ and $C = \bigcap_{x \in F} xS$. Because $F^{2n-1} \subset C$, we have $|S - C| < |F^{2n} - F^{2n-1}| = h_{2n} < \epsilon |A|$.

Let R be a projection maximal with the property that its translates xR (x in A) are mutually orthogonal. Since any non-zero projection dominates a non-zero projection whose translates under A are mutually orthogonal, it follows readily from the maximality of R that $I = A^{-1}AR$. We set $Q = AR$ and then, as in Lemma 4.1, set $P = (CQ)(I - (S - C)Q)$. By that lemma, $I - P$ is F-dispersed. On the other hand, since $\lambda(Q) \leqq 1/|A|$, we have $\lambda(I - P) \leqq \lambda((S - C)Q) \leqq |S - C| \lambda(Q) \leqq |S - C|/|A| < \epsilon$. We have proved that there exist F-dispersed projections of arbitrarily small measure, and the theorem follows from the remark following Proposition 4.1.

It follows that, if G is an abstract group containing a subset F satisfying the conditions of the theorem, then in any faithful representation as a freely-acting group of MP automorphisms of a hyperstonian measure space G is approximately finite. As an illustration, let G be the free product of two groups of order 2, and let F consist of the identity and the two generators of order 2. Then $h_1 = 3$ and $h_n = 2$ $(n > 1)$, and trivially, (4.1) holds.

LEMMA 4.2. *Let G be a freely-acting group which is the direct product of a finite subgroup D with an approximately finite subgroup K. Then G is approximately finite.*

Proof. Let F be a finite subset of K and let E_x be a determining function on K having finite support and such that $\lambda(E_x) > 1 - \epsilon/|D|$, for a pre-assigned $\epsilon > 0$. (That such exists follows from the results of Section 3.) Put $F'_x = \prod_{d \in D} dE_x$. Then F'_x is again a determining function on K with finite support, $\lambda(F'_x) > 1 - \epsilon$, for x in F, and furthermore, F'_x lies in the fixed algebra Z_D of D. If $g \in G$ has the (unique) representation $g = dk$ $(d \in D,\ k \in K)$, define $F_g = F'_k$. It is trivial to see that F_g is a determining function on G having finite support and satisfying $\lambda(F_g) > 1 - \epsilon$, for all g in DF. The result follows, since any finite subset of G is contained in a finite subset of the form DF.

With this we establish our main application of Theorem 1.

COROLLARY 4.1. *Any freely-acting abelian group is approximately finite.*

Proof. As we have noted above, it suffices to consider finitely generated abelian groups. By the familiar decomposition theorem, a finitely generated abelian group G is the direct product of a finite group and a free abelian group on (say) m generators. By Lemma 4.2, it suffices to assume that G is in fact free abelian $(m > 0)$. Let F consist of e, the m generators and their

inverses. By a simple inductive argument, one can show that $|F^n| = p_m(n)$, where p_m is a polynomial of degree m with rational coefficients. Therefore, $h_{2n}/(h_1 + \cdot\cdot\cdot + h_m) = [p_m(2n) - p_m(2n-1)]/p_m(n)$, and it is clear by elementary calculus that (4.1) holds. The corollary now follows from Theorem 1.

5. On the structure of full groups of type II. Let G be a full type II group of MP automorphisms of (M, λ), with fixed algebra denoted Z. Under the invariant metric $\delta(\alpha, \beta) = \lambda(I - F(\alpha, \beta))$, G is a topological group. Now in [3, Prop. 3.1] we proved that any full normal subgroup of G has the form G_C, for some (projection) C in Z. We shall require the following extension of this characterization.

PROPOSITION 5.1. *Any closed normal subgroup K of G has the form G_C, for some C in Z.*

Proof. It is easy to see that the full group $[K]$ generated by K will itself be normal in G, and therefore, by the result just cited, of the form G_C (C in Z). Clearly, we can assume in what follows that $C = I$, that is, $[K] = G$. The fixed algebra of K will now be Z and, for any projection P, its Z-carrier will be $\mathrm{LUB}_{k \in K}\, kP$. Our object is to prove $K = G$. First we prove

(5.1) Let C, D, R be mutually orthogonal non-zero projections with $E_Z(C) = E_Z(D) = E_Z(R)/2$. Then, given $\epsilon > 0$, there exist projections C_0, D_0, R_0 and a ρ in K with these properties: (i) $0 \neq C_0 \leqq C, D_0 \leqq D$, $R_0 \leqq R$ and $E_Z(C_0) = E_Z(D_0) = E_Z(R_0)/2$; (ii) $\rho(C_0) = D_0$, $F(e, \rho^2) \geqq C_0 + D_0$, $F(e, \rho) \geqq I - (C_0 + D_0 + R_0)$; (iii) $\lambda(C - C_0) < \epsilon$.

One has $R \leqq Z$-carrier $R = Z$-carrier $C = \mathrm{LUB}_{k \in K}\, kC$. So we can choose $0 \neq C' \leqq C$ and k in K such that $kC' \leqq R$. Using Maharam's lemma, dissect $C' = C_1 + (C' - C_1)$ so that $E_Z(C_1) = E_Z(C' - C_1)$, and choose $D_1 \leqq D$ with $E_Z(D_1) = E_Z(C_1)$. By [3, Lemma 3.2], there exists an α in G such that $\alpha(C' - C_1) = D_1$, $F(\alpha, e) \geqq I - [C' - C_1 + D_1]$, $\alpha^2 = \text{identity}$. Put $\beta = \alpha k \alpha$. Then $\beta \in K$ by normality of K, and

$$\beta(C_1 + D_1) = \alpha k C_1 + \alpha k (C' - C_1) = \alpha k C' = kC' \leqq R.$$

Similarly, choose a γ in G such that $\gamma C_1 = D_1$, $F(\gamma, e) \geqq I - (C_1 + D_1)$, $\gamma^2 = \text{identity}$. Put $\rho = \beta\gamma\beta^{-1}\gamma$. Again by normality, $\rho \in K$. Take $P \leqq C_1 + D_1$. We argue, $\rho P = \gamma P$; in fact, $\gamma P \leqq C_1 + D_1$ and

$$(C_1 + D_1)\beta(C_1 + D_1) \leqq (C_1 + D_1)R = 0,$$

so that $(C_1 + D_1)\beta^{-1}(C_1 + D_1) = 0$; γ is therefore the identity on $\beta^{-1}\gamma(C_1 + D_1)$,

and $\rho P=\beta\gamma(\beta^{-1}\gamma P)=\beta\beta^{-1}\gamma P=\gamma P$, as asserted. Set $R_1=\beta(C_1+D_1)$ and suppose $P\leqq I-(C_1+D_1+R_1)$. Then $\gamma P=P=\gamma\beta^{-1}P$, so $\rho P=P$. We have constructed projections C_1, D_1, R_1 and ρ in K which satisfy conditions (i) and (ii) of (5.1). The argument now proceeds by exhaustion. Let $(C_1{}^\alpha, D_1{}^\alpha, R_1{}^\alpha, \rho^\alpha)$ (α in $\mathfrak{a}$) be a maximal family with the properties (i) and (ii) and with the additional property that the $C_1{}^\alpha$ are mutually orthogonal (respectively, the $D_1{}^\alpha$ and the $R_1{}^\alpha$ are mutually orthogonal). One must have $C=\sum C_1{}^\alpha$, for otherwise we could apply the above construction to $C-\sum C_1{}^\alpha$, etc., to contradict maximality. Given $\epsilon>0$, choose a finite subset $\mathfrak{a}_0$ of $\mathfrak{a}$ such that $\lambda(C-\sum_{\alpha\in\mathfrak{a}_0} C_1{}^\alpha)<\epsilon$, and put $\rho=\prod_{\alpha\in\mathfrak{a}_0}\rho^\alpha$, $C_0=\sum_{\alpha\in\mathfrak{a}_0} C_1{}^\alpha$, etc. Then (C_1, D_1, R_1, ρ) satisfies (i) and (ii) and in addition $\lambda(C-C_0)<\epsilon$. This proves (5.1).

We turn to the proposition itself. Let α be an element of G having finite order. We will show that α can be approximated arbitrarily closely in the δ-metric by elements of K. It will follow that $\alpha\in K$, since K is closed. But elements of finite order are δ-dense in G ([3, Theorem 1]), and again since K is closed, we will have $K=G$.

Let Z_α be the (type I) fixed algebra of α. Since Z_α is of type II over Z, given $\delta>0$, we can choose a positive integer n with $6/(n+3)<\delta$ and mutually orthogonal projections $P_{-2}, P_{-1}, P_0,\cdots,P_n$ in Z_α with $E_Z(P_i)=1/(n+3)$. This done, for each $i=1,\cdots,n$, and $\epsilon=\delta/2n$, apply (5.1) to $R=P_{-2}+P_{-1}$, $C=P_i$, $D=P_0$. By (5.1), there exist ρ_i in K, $C_i\leqq P_i$, $D_i\leqq P_0$ such that $\rho_i(C_i)=D_i$, $\rho_i{}^2=$ identity on C_i+D_i, $\rho_i=$ identity on $P_1+\cdots+(P_i-C_i)+\cdots+P_n$, and $\lambda(P_i-C_i)<\epsilon$. For $i\geqq 1$, let $\alpha_i=P_i\alpha+(I-P_i)$ (well defined since P_i lies in Z), $\tau_i=\rho_i{}^{-1}\alpha_i{}^{-1}\rho_i\alpha_i$, and $\tau=\tau_1\cdots\tau_n$. By construction, $\tau\in K$. We claim that $\tau=\alpha$ on $\bigcup\alpha_i{}^{-1}C_i$. In fact, take $P\leqq\alpha_i{}^{-1}C_i$. Because $P\leqq P_i$, $\tau P=\tau_iP$. Because $\alpha_i{}^{-1}$ is the identity on P_0, $\tau_iP=\rho_i{}^{-1}\alpha_i{}^{-1}\rho_i(\alpha_iP)=\rho_i{}^{-1}\rho_i(\alpha_iP)=\alpha_iP=P$, verifying the claim. Now

$$\lambda\Big(\bigcup_{i=1}^{n}\alpha_i{}^{-1}C_i\Big)=\sum_{i=1}^{n}\lambda(C_i)\geqq\sum_{i=1}^{n}[\lambda(P_i)-\epsilon]=n/(n+3)-n\epsilon>1-\delta.$$

Therefore, $\delta(\alpha,\tau)<\delta$, and α lies in the closure of K. The proposition is proved.

Given a subgroup H of G, we denote by $H^\perp$ the centralizer of H in G, that is, the set of all elements in G which commute with each element of H.

COROLLARY 5.1. *Let H be a subgroup of the full type II group G. In order that H have the form G_C, for some C in the fixed algebra Z of G, it is necessary and sufficient that there exist a subgroup K of G such that $H=K^\perp$ and $G=HK$.*

Proof. Assuming that such a K exists, then H is closed in the topology of G, since any centralizer has this property. Also H is normal: if $\alpha = hk$ lies in G $(h \in H,\ k \in K)$, then $\alpha H \alpha^{-1} = hHh^{-1} = H$. Therefore, by Proposition 5.1, $H = G_C$ for some C in Z. Conversely, let $H = G_C$ $(C \in Z)$, and set $K = G_{(I-C)}$. Plainly $G = HK$. If $\alpha = hk$ $(h \in H, k \in K)$ lies in $K^{\perp}$, then k must lie in the center of K. But this forces $k = e$; K restricted to $(I - C)M$ is a full type *II* group, and it is easy to show that any full type *II* group has trivial center. Therefore, α lies in H.

It follows that local subgroups of G of the form G_C (C in Z_P) are characterized by the structure of G as an abstract group. We proceed now to show that general local subgroups G_P (P in M_P) can be similarly characterized. The basis for attack is the observation that any local subgroup of G is the centralizer of its centralizer; the technical problem is to sort out local subgroups from the class of all subgroups of G having this property.

LEMMA 5.1. *For any finite freely acting subgroup H of a full type II group G, $(H^{\perp})^{\perp} \subset [H]$.*

Proof. Choose a projection P such that the hP are mutually orthogonal and $\sum_{h \in H} hP = I$. Take ρ in $(H^{\perp})^{\perp}$. We will prove that $\bigcup_{h \in H} F(h, \rho) \geqq \rho^{-1}(P)$. Substitution of hP for P will then yield $\bigcup_{h \in H} F(h, \rho) = I$, and it will follow that $\rho \in [H]$. For any β in G_P, it is clear that $h\beta h^{-1} \in G_{hP}$, and therefore $\prod_{h \in H} h\beta h^{-1} \in H^{\perp}$. Fix h_0 in H and set $R = (\rho^{-1}P)h_0P$. Then for any $S \leqq R$, $\rho[\prod_h h\beta h^{-1}(S)] = \rho h_0 \beta h_0^{-1}(S)$, and $\prod_h h\beta h^{-1}\rho(S) = \beta\rho(S)$. Because $\rho \in (H^{\perp})^{\perp}$, this implies that $(\rho h_0)\beta = \beta(\rho h_0)$ on $h_0^{-1}R$. Now both $h_0^{-1}R \leqq P$ and $\rho R \leqq P$, and by appropriate choice of β, it is easy to see that $h_0^{-1}R = \rho R$. It now follows that ρh_0 commutes with all elements of $G_{h_0^{-1}R}$, and therefore is the identity on $h_0^{-1}R$. Thus, $h_0\rho$ is the identity on R, that is, $F(\rho, h_0^{-1}) \geqq (\rho^{-1}P)h_0P$. Therefore, $\bigcup_h F(\rho, h^{-1}) \geqq \sum_h (\rho^{-1}P)hP = \rho^{-1}P$, proving the lemma.

By the *carrier* of a non-vacuous subset A of G, we mean the projection $\mathrm{LUB}_{\alpha \in A}\ (I - F(\alpha, e))$ $(= \mathrm{GLB}[P \in M_P \mid G_P \supset A])$. The Z-carrier of A, that is, the Z-carrier of its carrier, is then $\mathrm{GLB}[C \in Z_P \mid G_C \supset A]$. In view of Corollary 5.1, the statement that two non-vacuous subsets of G have the same Z-carrier is purely algebraic (or in other words, depends only on the abstract group structure of G).

LEMMA 5.2. *Let G be a full type II group with fixed algebra Z. Let H and K be subgroups of G in the relation $H^{\perp} = K$, $K^{\perp} = H$, $H \cap K = \{e\}$.*

Assume that the fixed algebras Z_H and Z_K of H and K coincide with Z. Further, assume that any normal subgroup of H with the same Z-carrier as H has centralizer K. Then, $[H] \cap [K]$ is trivial.

Proof. Take $0 \neq C$ in Z_P and form $H_C = [Ch + (I-C) \mid h \in H]$. Evidently $H_C \subset K^{\perp} = H$. Likewise $K_C \subset K$. We will show that all the conditions imposed on H and K as subgroups of G remain valid for the subgroups H_C and K_C of the full type II group G_C on CM. To this end, suppose $\alpha \in G_C$ lies in $H^{\perp}$. Then $\alpha = C\alpha + (I-C)$ lies in $H^{\perp} = K$, and hence $\alpha \in K_C$. Likewise, the centralizer of K_C in G_C is H_C, and obviously, $H_C \cap K_C$ is trivial. The fixed algebra of H_C or K_C coincides with that of G_C (viz. $CZ + (I-C)M$). Finally, suppose that N is a normal subgroup of H_C with the same Z-carrier as H_C. Then $NH_{(I-C)}$ is a normal subgroup of H with the same Z-carrier as H. Therefore, its centralizer in G is $K = K_C K_{(I-C)}$. It follows that the centralizer of N in G_C is K_C. Therefore, our hypotheses are invariant under restriction to Z-summands.

We prove the lemma by an indirect argument: assume that, for certain elements $h_0 \in H$, $k_0 \in K$, one has $(I - F(h_0, e))F(h_0, k_0) \neq 0$ (so that $[H] \cap [K]$ is non-trivial). We shall obtain a contradiction.

To begin, note that for h_1, h_2 in H, $F(h_1, h_2) \in Z_K = Z$. Therefore, $I - F(h_0, e)$ is a non-zero projection in Z. Dropping to a Z-summand, we can assume $F(h_0, e) = 0$, retaining the other hypotheses; and in the same way, we can assume that Z-carrier $F(h_0, k_0) = I$.

Let A be a subset of the conjugate class $C(h_0)$ of h_0 in H which contains h_0, whose elements are independent, meaning $F(x, y) = 0$ if $x \neq y, x, y$ in A, and which is maximal with these properties. How if $x, y \in A$, then $F(x, k_0)F(y, k_0) \leqq F(x, y) = 0$; and because of the relation $sF(h, k) = F(shs^{-1}, k)$ (valid for s, h in H, k in K), we see that $\lambda(F(x, k_0)) = \lambda(F(y, k\ ))$ $(x, y$ in $A)$. Therefore, A is finite. For z in $C(h_0)$, define $E(z) = I - \sum_{x \in A} F(x, z)$. Clearly, $E(z) \in Z_P$. We claim that $\mathrm{LUB}_{z \in C(h_0)} E(z) \neq I$. Otherwise, there exist z_n in $C(h_0)$ and C_n in Z_P, $C_n \leqq E(z_n)$, such that $\sum C_n = I$. If $z_n = s_n h_0 s_n^{-1}$, let $s = \sum C_n s_n$. This element s lies in H; H is closed in the topology of G, being a centralizer. In turn, let $z = sh_0s^{-1}$. Then for arbitrary x in A, $F(x, z) = \sum_n C_n F(x, z_n) \leqq \sum_n E(z_n)(I - E(z_n)) = 0$. This contradicts the maximality of A. Therefore, there exists $0 \neq C$ in Z_P such that $\sum_{x \in A} F(x, z) \geqq C$, for each z in $C(h_0)$. Dropping to a Z-summand, therefore, we can assume $\sum_{x \in A} F(x, z) = I$, for each z in $C(h_0)$.

Let B be the boolean algebra generated by Z and the projections $F(x, k_0)$ $(x$ in $A)$. If $z \in C(h_0)$, then

$$F(z,k_0)=\sum_{x\in A}F(z,k_0)F(x,z)=\sum_{x\in A}F(x,z)F(x,k_0)\in B.$$

In particular, then, B is H-invariant. Further, since

$$I=Z\text{-carrier }F(h_0,k_0)=\mathrm{LUB}_{s\in H}\,sF(h_0,k_0)=\mathrm{LUB}_{z\in C(h_0)}F(z,k_0),$$

one has $\sum_{x\in A}F(x,k_0)=I$. Therefore, if $n=|A|$, B is of type I_n over Z. We claim that the representation of H obtained by restricting H to act on B is faithful. For this, suppose that s in H determines the identity automorphism of B. Then for each x in A,

$$F(x,k_0)=sF(x,k_0)=F(sxs^{-1},k_0)=\sum_{y\in A}F(y,sxs^{-1})F(y,k_0),$$

and multiplication by $F(x,k_0)$ gives $F(x,k_0)=F(x,sxs^{-1})F(x,k_0)$. Therefore, $F(x,sxs^{-1})\geqq Z\text{-carrier }F(x,k_0)=I$, proving that s commutes with all elements of A. For any z in $C(h_0)$, x in A, $F(x,szs^{-1})=sF(s^{-1}xs,z)=F(x,z)$, so that

$$F(szs^{-1},z)=\sum_{x\in A}F(szs^{-1},z)F(z,x)=\sum_{x\in A}F(szs^{-1},x)F(z,x)=\sum_{x}F(z,x)=I,$$

and s commutes with z. It follows that s lies in the centralizer of the normal subgroup $N(h_0)$ of H generated by h_0. By construction (viz. $F(h_0,e)=0$), this subgroup has carrier I. By assumptions of the lemma, therefore, $N^{\perp}=K$, and $s\in H\cap K=\{e\}$. This proves that the restriction of H to B is faithful.

Let S be the subgroup of H generated by A. Restricted to B, S is finite ([3, Lemma 4.4]). Therefore, S is finite. If we form the finite boolean algebra generated by the projections $F(s_1,s_2)$ in Z (s_1, s_2 ranging over S), then the restriction of S to an atom of this algebra is freely acting (after identification of elements having the same action on this atom). Dropping to a Z-summand again, we can suppose S is freely-acting. Now the relation $\sum_{x\in A}F(x,z)=I$ (x in $C(h_0)$) gives $z=\sum_{x}F(x,z)x$. It follows in turn that each element in the normal subgroup $N(h_0)$ of H generated by h_0 has a representation $\sum_{y\in S}C_y y$ (C_y in Z). On this count, S and $N(h_0)$ have the same centralizer, viz. K. Therefore, $(S^{\perp})^{\perp}=H$, and since $Z_H=Z$, it follows that $(S^{\perp})^{\perp}$ is a group of type II. On the other hand, by Lemma 5.1, $(S^{\perp})^{\perp}$ is a subgroup of a type I group, and is therefore of type I. This contradiction establishes the lemma.

Now let $H=G_P$ be a local subgroup of the full type II group G. Since the center of H must be trivial, it follows that $H^{\perp}=G_{(I-P)}$. Therefore,

(5.2) If $H^{\perp}=K$, then $K^{\perp}=H$ and $H\cap K=\{e\}$.

H and K also satisfy

(5.3) Any normal subgroup N of H (respectively, of K) with the same Z-carrier as H (respectively, as K) has centralizer K (respectively, H).

This follows since, by Proposition 5.1, the closure N^- of N must coincide with H, and since, trivially, N and N^- have the same centralizer. Finally,

(5.4) If α is an element of G not in HK, then there exists an element $h \neq e$ in H (respectively, in K) such that $\alpha h \alpha^{-1}$ lies in K (respectively, in H).

In fact, if α is not in HK, then $P \neq \alpha(P)$, and so $R = P\alpha^{-1}(I-P) \neq 0$. G_R is contained in H, and if h is an element of G_R not e, then $\alpha h \alpha^{-1} \in G_{\alpha(R)}$ and $G_{\alpha(R)} \subset K$.

PROPOSITION 5.2. *The conditions* (5.2), (5.3), *and* (5.5) *characterize local subgroups H of a full type II group G.*

Proof. Our task is to show that, if H (and $K = H^\perp$) are subgroups of G satisfying these three conditions, then H is local. Write Z_H, Z_K, and Z_{HK} $(= Z_H \cap Z_K)$ for the fixed algebras of H, K, and HK. Let P_H and P_K be the carriers of H and K. Our assumptions entail that $P_H \cup P_K = \mathrm{I}$; for any α in $G_{I-P_H \cup P_K}$ lies in both $H^\perp$ and $K^\perp$, that is, by (5.2), in $K \cap H = \{e\}$. Also, both P_H and P_K lies in Z_{HK}; this follows readily from the definition of carrier and the fact that $H^\perp = K$, $K^\perp = H$. Now let C be any projection in Z_{HK} satisfying $C \leqq P_H$, let $\bar{C}$ denote the Z-carrier of C, and write

$$N = [(I - \bar{C} + C)h + \bar{C} - C \mid h \in H].$$

Each element of N lies in $K^\perp = H$ and N is evidently normal in H. The carrier of N is $(I - \bar{C} + C)P_H = (I - \bar{C})P_H + C$, so that the Z-carrier of N is $(I - \bar{C})P_H + \overline{CP_H} = P_H$. Therefore, by condition (3.2), $N^\perp = K$. Now if $h \in H$, then $(\bar{C} - C)h + (I - \bar{C} + C)$ lies in $N^\perp$, and therefore, because $H \cap K = \{e\}$, $(\bar{C} - C)P_H = 0$. This shows that $Z_{HK}P_H = ZP_H$. Likewise, $Z_{HK}P_K = ZP_K$. It follows that Z_{HK} is the boolean algebra generated by adjunction of P_H and P_K to Z. As in Lemma 5.2, it follows readily that, for each projection C in Z_{HK}, the three conditions (5.2)-(5.4) are valid for the subgroups H_C and K_C of G_C on CM.

We will prove that $P_H P_K = 0$. Using this, it then follows that $G_{P_H} = G_{I-P_K} \subset K^\perp = H \subset G_{P_H}$, namely, $H = G_{P_H}$ is local, and the proposition will be proved. As the basis of an indirect proof, we assume $P_H P_K \neq 0$. Dropping to $G_{P_H P_K}$ on $P_H P_K M$, we retain conditions (5.2)-(5.4), as noted above. Therefore, for the sake of notation, we may assume $P_H = P_K = I$. We then have $Z_{HK} = Z$.

Next, we show that $Z_H = Z_K = Z$. In fact, take $0 \neq P \in Z_H$. We will

show that P is K-invariant. This will prove that $P \in Z_{HK} = Z$, showing that $Z_H = Z$; similarly, $Z = Z_K$ will follow. Define $N = [Ph + I - P \mid h \text{ in } H]$. N is a subgroup of G with carrier P, and $hNh^{-1} = N$ for each h in H. If we knew that N were a normal subgroup of H, then for each k in K, $Ph + I - P = k(Ph + I - P)k^{-1} = (kP)h + I - kP$, so kNk^{-1} has carrier kP; but since the carrier of any normal subgroup of H is readily seen to be K-invariant, we would have $kP = P$. Proceeding indirectly, assume that N is not a subgroup of H. Then, an h_0 in H exists such that $\alpha = P h_0 + I - P$ does not lie in H. We consider two cases. Case (i): α does not lie in HK. In this event, by (5.4), there exists $h \neq e$ in H such that $\alpha h \alpha^{-1} \in K$. Now $\alpha h \alpha^{-1} = Ph_0 h h_0^{-1} + (I - P)h$. Because $\alpha h \alpha^{-1}$ commutes with all elements of H, and in particular with h_0, we get

$$Ph_0 h h_0^{-1} + (I - P)h = Ph + (I - P)h_0^{-1} h h_0,$$

so that $Ph_0 h h_0^{-1} = Ph$, and $\alpha h \alpha^{-1} = Ph + (I - P)h = h \in H$. Therefore, $\alpha h \alpha^{-1}$ lies in $H \cap K = \{e\}$, a contradiction. Case (ii): α lies in HK not H. Thus $\alpha = hk$ $(k \neq e)$. Let L be the normal subgroup of K generated by k. We note that each element n of L leaves P fixed. In fact, it suffices to verify this for $n = uku^{-1}$ $(u \in K)$, and one has $uku^{-1}P = uhku^{-1}(h^{-1}P) = u\alpha u^{-1}P = u[Ph_0 u^{-1}P + (I - P)u^{-1}P] = (uP)P + (I - uP)P = P$. It follows that each n in L commutes with α. The carrier P_L of L lies in $Z_H \cap Z_K = Z$. The subgroup $LK_{(I-P_L)}$ of K has the same Z-carrier (namely, I) as K, and so by (5.3), its centralizer is H. But $P_L\alpha + I - P_L \in (LK_{I-P_L})^{\perp} = H$, and $(I - P_L)\alpha + P_L = (I - P_L)h + P_L \in H$, whence

$$\alpha = (P_L + (I - P_L)\alpha)(P_L\alpha + I - P_L)$$

lies in H, a contradiction. Thus we have proved that Z_H (and similarly, Z_K) coincides with Z.

We can now apply Lemma 3.2 to conclude that $[H] \cap [K] = \{e\}$. Clearly, we can choose α in $[H]$ so that $F(\alpha, e)$ does not lie in Z. Then α does not lie in H (for h in H implies $F(h, e) \in E_K = Z$), and also α does not lie in HK; for if $\alpha = hk$, then $k = h^{-1} \in$ lies in $[H] \cap [K] = \{e\}$. We can apply (5.4): there exists an $h \neq e$ in H such that $\alpha h \alpha^{-1} \in K$. Because $\alpha h \alpha^{-1} \in [H]$, we have $\alpha h \alpha^{-1} \in K \cap [H] = \{e\}$, a contradiction. Therefore, the assumption that $P_H P_K \neq 0$ is untenable, and the proposition follows.

In view, therefore, of Corollary 5.1 and Proposition 5.2, the conditions (5.2)-(5.4) provide a purely algebraic characterization of local subgroups of G. It is merely the existence of this characterization, rather than its specific detail, which is basic in the following

THEOREM 2. *Any group isomorphism between two full groups of type II is implemented by a weak equivalence of these groups.*

Proof. Let G and G' be full type II groups on (M,λ) and (M',λ'), and let ϕ be an isomorphic mapping of G on G'. In view of Corollary 5.1, for each projection C in Z_G there exists a uniquely determined projection $\theta(C)$ in $Z_{G'}$ such that $\phi(G_C) = G'_{\theta(C)}$. Application of the same remark to ϕ^{-1} shows that θ is a 1:1 mapping of Z_G on $Z_{G'}$. Furthermore, θ conserves products, for $G'_{\theta(CD)} = \phi(G_{CD}) = \phi(G_C \cap G_D) = G'_{\theta(C)} \cap G'_{\theta(D)} = G'_{\theta(C)\theta(D)}$. It follows that θ is a boolean isomorphism—a 1:1 mapping between boolean algebras is a boolean isomorphism if and only if it conserves products —. In turn, it follows that θ induces a $*$-isomorphism of Z_G on $Z_{G'}$. In particular, then, two subgroups in G have the same Z_G-carrier if and only if the corresponding subgroups in G' have the same $Z_{G'}$-carrier. Let $H = G_P$ be an arbitrary local subgroup of G. The subgroup $H' = \phi(H)$ of G' will then satisfy the conditions (5.2)-(5.4). By Proposition 5.2, H' will be local, and must therefore have the form $H' = G'_{\theta(P)}$, for some projection $\theta(P)$ in M'. Now apply the argument given above for the pair Z_G, $Z_{G'}$ to the pair M, M'. It follows that θ induces a $*$-isomorphism of M on M'. Now

$$G'_{\theta(xP)} = \phi(G_{xP}) = \phi(xG_Px^{-1}) = \phi(x)G'_{\theta(P)}\phi(x)^{-1} = G'_{\phi(x)\theta(P)},$$

for all x in G. This shows that $\theta(xP) = \phi(x)\theta(P)$, or in other words, $\phi(x) = \theta x\theta^{-1}$. This proves that ϕ is implemented by θ, and the theorem follows.

6. Appendix: connections with von Neumann algebras. Let M be a von Neumann algebra with a faithful normal (numerical) trace Tr. Let H be a maximal abelian self-adjoint ($=$ MASA) subalgebra of M having no minimal non-zero projections, and let $N(H)$ denote the normalizer of H, that is, the collection of all unitary operators U in M such that $UHU^{-1} = H$. In the foregoing terminology, the pair (H, Tr) is a non-atomic hyperstonian measure space, and the mapping $U \to \phi_U$, where $\phi_U(A) = UAU^{-1}$, is a homomorphism from $N(H)$ into the group of all MP automorphisms of (H, Tr), its kernel being the group H_U of all unitaries in H, this by the maximality of H. Let us denote by G the group of all such MP automorphisms ϕ_U $(U \in N(H))$. It is trivial that G is a full group; in fact, the most general element of $[G]$ must have the form $\sum Q_n\phi_{U_n}$, the Q_n (respectively, the $\phi_{U_n^{-1}}(Q_n)$) being mutually orthogonal projections in H with

$$\sum Q_n = I \quad (\sum \phi_{U_n^{-1}}(Q_n) = I),$$

and such an automorphism is already implemented by the element $V = \sum Q_n U_n$ of $N(H)$ (the $\sum$ being taken in the strong operator topology). Similarly, if N is a von Neumann subalgebra of M containing H—or as we shall say, an intermediate subalgebra—then the group K of all ϕ_U, U ranging over $N(H) \cap N$, is a full subgroup of G. In operator theory, various concepts of type are attached to M and its intermediate subalgebras; similarly, in the foregoing study of groups of MP automorphisms, we have attached concepts of type to G and its subgroups. Our object now is to compare these terminologies in the special case when the MASA subalgebra H is *regular* (Dixmier, [2]), that is, when the von Neumann subalgebra $R(N(H))$ of M generated by $N(H)$ is all of M, and to indicate (without elaboration) how some of our theorems on MP groups can be carried over to von Neumann algebras.

Technically, the discussion to follow makes heavy use of the concept of conditional expectation in a (finite) von Neumann algebra, and we review this matter briefly (see Umegaki, [5]). With M, Tr as above, let N be an arbitrary von Neumann subalgebra of M containing the center Z of M. Then, for each A in M, there exists a uniquely-determined element $E_N(A)$ of N such that $\mathrm{Tr}(E_N(A)B) = \mathrm{Tr}(AB)$, for all B in N. $E_N(A)$ is called the conditional expectation of A relative to N. (The existence of $E_N(A)$ follows readily from a Radon-Nikodym theorem, which shows also that $E_N(A)$ is independent of the particular faithful normal trace used on M.) The key properties of E_N follow from the defining formula: $A \rightarrow E_N(A)$ is a positive *-linear mapping, conserving the identity I, and satisfying

(6.1) $\quad E_N(AB) = E_N(A)B$, for all A in M, B in N, and

(6.2) $\quad E_N(A^*A) \geqq E_N(A)^*E_N(A)$, for all A.

If $[A] = (\mathrm{Tr}(A^*A))^{\frac{1}{2}}$ denotes the trace norm on M, then (6.2) implies that $[E_N(A)] \leqq [A]$. When N is the center of M, then E_N is simply the center-valued trace on M.

LEMMA 6.1. *Let H be a regular MASA subalgebra of M, let N be an intermediate von Neumann subalgebra of M, and let K be the full group of MP automorphisms of (H, Tr) induced by $N(H) \cap N$. Then H is regular as a MASA subalgebra of N, and for each U in $N(H)$, there exists a W in $N(H) \cap N$ such that $E_N(U) = E(K, \phi_U)W$.*

Proof. In view of (6.1) and the fact that $H \subset N$, we have

$$E_N(U)A = \phi_U(A)E_N(U), \tag{6.3}$$

for all A in H. Let $S = [E_N(U)^*E_N(U)]^{\frac{1}{2}}$. Using (6.3) and the fact that $E_N(U)^* = E_N(U^{-1})$, we have $S^2A = AS^2$, for all A in H. By maximality

of H, this forces $S^2 \in H$. Consequently, $S \in H$. Let $E_N(U) = VS$ be the polar decomposition of $E_N(U)$. Here V is a partial isometry with initial projection $E = V^*V \in H$ and terminal projection $F = VV^*$. Substitution of VS for $E_N(U)$ in (6.3) yields $\phi_U(A)VS = VSA + VAS$, and from this and the fact that $VE = V$, we deduce

$$\phi_U(A) = VAV^*, \tag{6.4}$$

for all A in H. Substitution of E for A in (6.4) shows immediately that $\phi_U(E) = F$. In turn, we deduce readily from (6.4) that $V^*UA = AV^*U$, for all A in H. This implies that $V^*U \in H$, and so $V^*U = E_N(V^*U) = V^*E_N(U) = S$. Thus V^*U is a non-negative self-adjoint partial isometry in H, and consequently, a projection. Therefore, $S = E$, and $E_N(U) = V$. We note that $VHV^* \subset H$, by (6.4), and also that $V^*HV \subset H$, by the same argument applied to U^{-1}.

We claim that we can extend V to a unitary W in $N(H) \cap N$. For this purpose, by Zorn, it evidently suffices to prove the following: if V is a partial isometry in N, with initial and terminal projections in H, and with $VHV^* \subset H$, $V^*HV \subset H$, then if V admits no proper extension satisfying these conditions, it must already be unitary. Suppose to the contrary that $V^*V = E \neq I$. Then, by the finiteness of N, $VV^* = F \neq I$, and $I - E$ and $I - F$ are equivalent mod N. So there exists a partial isometry T in N such that $T^*T = I - E$ and $TT^* = I - F$. Making our first use of regularity of H, given $\epsilon > 0$, we can find elements $U_1, \cdots, U_n$ of $N(H)$ and scalars $a_1, \cdots, a_n$ such that $[T - \sum_{i=1}^{n} a_i U_i] < \epsilon$. Since E_N is norm-depressing, the same inequality holds with $E_N(U_i)$ replacing U_i. Using the first part of the proof, let V_i be the partial isometries in N such that $E_N(U_i) = V_i$. If we choose $\epsilon < [T]$ and observe that $[T - \sum a_i(I-F)V_i(I-E)] < \epsilon$, we deduce that some $V' = (I-F)V_i(I-E)$ is $\neq 0$. But V' is a partial isometry (since H is abelian) with initial projection $\leqq I - E$, terminal projection $\leqq I - F$, such that $V'HV'^* \subset H$, $V'^*HV' \subset H$. It follows that $V + V'$ is a partial isometry properly extending V and having the same properties. This contradiction establishes our claim.

We have proved that $E_N(U)$ has the form FW, for some projection F in H and some W in $N(H) \cap N$. It follows immediately that $F\phi_U(A) = F\phi_W(A)$, for all A in H, so that $F \leqq E(K, \phi_U)$. On the other hand, writing $E_1 = E(K, \phi_U)$, we know by [3, Lemma 3.4] that there exists a ϕ_W in K such that $E_1\phi_W(A) = E_1\phi_U(A)$, for all A in H. From this relation and the maximality of H, we deduce that $E_1UW^{-1} \in H$, and in turn, that $E_1U = E_1W_1$, for some W_1 in $N(H) \cap N$. Now

$$(I-F)E_1U=(I-F)E_N(E_1U)=0,$$

and this shows that $(I-F)E_1=0$, or $E_1\leqq F$. Accordingly, $F=E(K,\phi_U)$.

It remains only to prove that H is a regular MASA subalgebra of N. To this end, let N_1 denote the von Neumann subalgebra of N generated by $N(H)\cap N$. The characterization of $E_N(U)$ just obtained shows that, for each U in $N(H)$, $E_N(U)\in N_1$. This and the regularity of H in M force $E_N(A)\in N_1$, for all A in M. Obviously, therefore, $N=N_1$, and the lemma is established.

Application of the lemma in the case $N=H$ yields the following criterion: in order the ϕ_U be freely-acting, it is necessary and sufficient that $E_H(U)=0$.

PROPOSITION 6.1. *There is a* $1:1$ *correspondence between full subgroups* K *of* G *and intermediate von Neumann subalgebras* $H\subset N\subset M$ *of* M, *obtained by associating with each full subgroup* K *the intermediate subalgebra* $R[U\mid\phi_U\in K]$ *and with each intermediate subalgebra* N *the subgroup* $[\phi_U\mid U\in N(H)\cap N]$. *This correspondence conserves type: if* N *is of type I (respectively type II) as an operator algebra, then the corresponding subgroup is of type I (respectively, type II) as a group of MP automorphisms of* (H,Tr).

Proof. We have already remarked that the subgroup associated with an intermediate subalgebra N is full. Because H is regular in N, by Lemma 6.1, the subalgebra associated in turn with this subgroup is the original N.

On the other hand, consider a full subgroup K of G, and form the corresponding subalgebra $N=R[U\mid\phi_U\in K]$. We must show that, if $V\in N(H)\cap N$, then $\phi_V\in K$. To see this, write $F=\operatorname{LUB}_{\alpha\in K}F(\phi_V,\alpha)$, and choose a ϕ_U in K such that $F(\phi_V,\phi_U)=F$. Then $\phi_{U^{-1}V}$ leaves F absolutely fixed, and furthermore, for each ϕ_W in K, $\phi_{W^{-1}V}$ is freely-acting on $I-F$. By Lemma 6.1, this entails $E_H(W^{-1}V)(I-F)=0$. We want to establish that $F=I$. Given $\epsilon>0$, we can choose (by Definition of N) unitaries $U_1,\cdots,U_n$ with $\phi_{U_i}\in K$ and scalars $a_1,\cdots,a_n$ such that $[V-\sum a_iU_i]<\epsilon$. It follows that $[I-F-\sum a_iV^{-1}U_i(I-F)]<\epsilon$, or on application of E_H, that $[I-F-\sum a_iE_H(V^{-1}U_i)(I-F)]<\epsilon$. Because each $E_H(V^{-1}U_i)(I-F)=0$, this gives $[I-F]<\epsilon$, and it follows that $F=I$. This establishes the existence of a $1:1$ correspondence.

Now let N be an intermediate subalgebra, and K the corresponding subgroup of G. It is trivial that the center of N coincides with the fixed algebra Z_K of K. Let P be a non-zero projection in H which is abelian in N in the sense of operator theory—that is, $PNP\subset PZ_K$—. Then a fortiori $PH\subset PZ_K$,

and P is abelian over Z_K in the measure-theoretic sense. Conversely, if P is an abelian projection over Z_K in the measure theoretic sense, then we claim that P is abelian in N in the operator sense. For this, it suffices to show that $PUP \in PZ_K$, for each U in $N(H) \cap N$. Let $F = F(\phi_U, e)$. By Lemma 6.1, $E_H(U) = FU$ $(= UF)$. Furthermore, by the measure-theoretic definition of abelian projections, $P\phi_U(P)E_H(U) \subset PZ_K$, and $P(I-F)\phi_U(P(I-F)) = 0$ ([3, Lemma 4.1]). It follows that $PUP = P\phi_U(P)U = P\phi_U(P)FU = P\phi_U(U) \in PZ_K$, proving that P is abelian in N in the operator sense.

Assume that N is of type I as a von Neumann algebra—namely, that each non-zero projection dominates an abelian projection —. We claim that K is of type I. If not, there exists a non-zero projection C in Z_K such that K_C is of type II on CH. Choose an abelian projection P in N with central carrier C. By Maharam's lemma, since CH is of type II over CZ_K, there exists a projection Q in H with $E_{Z_K}(Q) = E_{Z_K}(P)$. This means that P and Q have the same center-valued trace in N, and that they are therefore equivalent mod N. Since equivalence preserves abelian projections, Q is therefore an abelian projection $\leqq C$, a contradiction.

Assume that N is of type II, in the sense that it contains no abelian projections. Then the corresponding K must be of type II, for otherwise H would contain a projection abelian over Z_K. The proposition is proved.

Assume now that M is type II. Then we shall call M approximately finite if, for each $\epsilon > 0$ and each finite set $X_1, \cdots, X_n$ of operators in M, there exists a type I subalgebra N of M having the same center Z as M and such that $[X_i - E_N(X_i)] < \epsilon$, for all i. (In the terminology of Widom [6], this is approximate finiteness (A-2).)

COROLLARY 6.1. *If G is approximately finite, then so is M.*

Proof. By the regularity of H, it suffices to consider the case $X_i \in N(H)$, $i = 1, \cdots, n$. Because G is approximately finite and of type II, given $\epsilon > 0$, we can choose a finite freely-acting subgroup K of G such that $[I - E([K], \phi_U)] < \epsilon$, for all i. Denote by N the intermediate subalgebra corresponding to K. By Lemma 6.1, we have $[X_i - E_N(X_i)] = [I - E([K], \phi_{X_i})]$. Therefore, we can achieve the desired approximation in a type I intermediate subalgebra. It remains of course to achieve the approximation in a type I subalgebra having the same center Z as M, and for this, the following assertion contains the main technical step.

(6.5) Let K be a finite freely-acting subgroup of G, let N be the intermediate subalgebra corresponding to $[K]$, and let $Y_1, \cdots, Y_n$ be arbitrary elements of N. Then, for each $\epsilon > 0$ and each non-zero central projection D,

there exists a non-zero central projection $C \leqq D$ and a type I subalgebra N_1 of the relative ring M_C having the same center Z_C as M_C and such that $[Y_iC - E_{N_1}(Y_i)]^2 < \epsilon^2 \operatorname{Tr}(C)$.

For proof, we shall need the following elementary result concerning group extensions. Let ϕ be a homomorphism with abelian kernel mapping a group A onto a finite group Q. Then any finite subgroup C of A which is mapped isomorphically into Q by ϕ can be imbedded in a finite subgroup B of A which is mapped isomorphically onto Q.

Applying this result in our present situation with $Q = K$, $A = \phi^{-1}(K)$, and C trivial, we can chose a finite group of unitaries U_k in $N(H)$ such that $\phi_{U_k} = k$, for each k in K. Each element of the intermediate subalgebra then has a unique representation $\sum A_kU_k$ (A_k in H). In proving (6.5), then, it suffices to consider the case $Y_i = A_iU_{k_i}$ ($A_i \in H, k_i \in K$).

By spectral theory, for each i, H contains a finite set S_i of mutually orthogonal projections such that, for an appropriate linear combination A'_i of the projections in S_i, $\| A_i - A'_i \| < \epsilon/2$ ($\|$ denoting the uniform norm). Choose a projection P in H such that the kP are mutually orthogonal and $\sum kP = I$, and let $\mathfrak{B}$ be the smallest K-invariant Boolean algebra containing the projection P and all the sets S_i. $\mathfrak{B}$ is then finite, and each of its atoms is abelian over the fixed algebra Z_k of K. Choose an atom from each K-orbit in the set of atoms, denoting the atoms so chosen as $P_1, \cdots, P_m$. Thus, $I = \sum_{i=1}^{m} \sum_{k} kP_i$. There exist disjoint central projections $C_1, \cdots, C_r$ with $\sum C_j = I$ so that, for each j, the non-zero projections in the set $C_jP_1, \cdots, C_jP_m$ all have central carrier C_j. Choose j such that $DC_j \neq 0$, and then choose a non-zero central projection $C \leqq DC_j$ such that, if $R_1, \cdots, R_s$ denote the non-zero terms among the CP_i, then all the R_i have central carrier C and $E_Z(R_1) \wedge \cdots \wedge E_Z(R_s)$ is bounded away from 0 on C. Take $\delta > 0$, δ to be specified later. Then, by application of Maharam's lemma, there exists an integer t and mutually orthogonal projections $Q_0, \cdots, Q_{t-1}$ in H such that $\sum Q_i = R_1 + \cdots + R_s$, all the Q_i have the same expectation relative to Z, and for each R_j, the sum R'_j of those $Q_i \leqq R_j$ satisfies $\operatorname{Tr}(R_j - R'_j) < \delta$. Now, δ is to be chosen so small that an appiopriate linear combination A''_i of the Q_i will satisfy $[A''_iC - A'_iC] < \epsilon(\operatorname{Tr}(C))^{\frac{1}{2}}/2$, for all i. Next, choose an automorphism τ in G_C which commutes with all elements of K and satisfies $\tau(Q_i) = Q_{i+1}$ (indices mod t), $\tau^t =$ identity. Let L be the freely-acting finite subgroup of G_C generated by τ and K_C. Applying the previously-cited result on group extensions, we can choose a finite group of unitaries U_x in $N(H)_C$ containing the U_k and such that $\phi_{U_x} = x$, for each x in L. Let A be the

von Neumann subalgebra of M consisting of all operators of the form $\sum_{x \in L} C_x x Q_0$ (C_x in Z_C). Because the projections xQ_0 are mutually orthogonal and $\sum xQ_0 = C$, it is easy to see that the L-invariant elements of A are precisely elements of Z_C. In turn, let N_1 be the subalgebra of M_C consisting of all operators of the form $\sum_{x \in L} A_x U_x$ (A_x in A). Then the center of N is precisely Z_C. Furthermore, it is easy to see that the elements $A''_i U_{k_i} C$ of N satisfy the inequality $[A_i U_{k_i} C - A''_i U_{k_i} C] < \epsilon(\mathrm{Tr}(C))^{\frac{1}{2}}$. (6.5) follows directly.

By exhaustion, we can express the identity I as a sum of mutually orthogonal central projections C satisfying the conclusion of (6.5). If N_2 denotes the direct sum of the corresponding N_1's, then N_2 is a type I subalgebra of M having center Z and such that $[Y_i - E_{N_2}(Y_i)] < \epsilon$, for all i. Applying this result to the operators $Y_i = E_N(X_i)$, we obtain $[X_i - E_{N_2}(X_i)] < 3\epsilon$, and the corollary follows.

Conjecturally, the converse to this corollary—that the group G attached to an arbitrary regular MASA in an approximately finite von Neumann algebra M is approximately finite—is false.

Remark 6.1. To conclude, we shall discuss a case, involving the so-called group-measure space examples of von Neumann algebras, in which application of the foregoing theory to von Neumann algebras is relatively straightforward.

Given a freely acting group G of MP automorphisms of an (abstract) hyperstonian measure space (H, λ), there exists a canonical procedure for constructing a finite von Neumann algebra—see for example [1, p. 132]—. To review details briefly, one forms the hilbert algebra $\mathfrak{A}$ consisting of all functions $A(g)$ from G to H, each having finite support, with the product $(AB)(g) = \sum_h A(h) h[B(h^{-1}g)]$, the adjoint $(A(\cdot))^*(g) = g(A(g^{-1})^*)$, and the inner product $(A, B) = \sum_h \lambda(A(h)B(h)^*)$. The left ring $\mathcal{L}$ of $\mathfrak{A}$ is then a von Neumann algebra with a distinguished faithful normal numerical trace Tr, a distinguished regular MASA subalgebra H_1, and a distinguished group G_1 of unitaries in $N(H_1)$ which, with H_1, generates $\mathcal{L}$, these having the following property: there exists an isomorphism θ of H on H_1 such that i) $\mathrm{Tr}(\theta(A)) = \lambda(A)$, for all A in H, and ii) the group of MP automorphisms of (H_1, Tr) induced by G_1 is precisely the group $\theta G \theta^{-1}$. The full group $[G]$ corresponds under θ to the full group determined by $N(H_1)$.

Carrying this one step further, suppose that M is a von Neumann algebra with a faithful normal trace Tr, that H is a *MASA* in M, and that $N(H)$ contains a group of unitaries U_g which, with H, generates M, and which

induces a freely-acting group of MP automorphisms of (H, Tr). One can apply the above construction to obtain a new von Neumann algebra $\mathcal{L}$. Now, we assert, $\mathcal{L}$ is isomorphic to M. To see this, let M_0 be the dense *-subalgebra of M consisting of all operators $\sum_g A_g U_g$, where $A_g \in H$ and $=0$ except for finitely many g's. Under the inner product $(A,B) = \mathrm{Tr}(B^*A)$, M_0 is a hilbert algebra, and M is isomorphic to the left ring of this hilbert algebra. We define a mapping from M_0 to the hilbert algebra $\mathfrak{A}$ by sending $A \in M_0$ to the element $E_H(AU_{g^{-1}})$ of $\mathfrak{A}$. Using the fact that $E_H(U_g) = 0$ unless $g = e$, it is easy to verify that T is a hilbert algebra isomorphism, namely, that it conserves products, adjoints, and norm, and is onto. Therefore, T induces an isomorphism of the left rings of these two hilbert algebras, that is, of M and $\mathcal{L}$. In summary, then, in order that a von Neumann algebra with a faithful normal trace arise from this group-measure space construction, it is necessary and sufficient that this algebra contain a MASA subalgebra and a group of unitaries in the normalizer of this subalgebra satisfying the conditions set down at the beginning of this paragraph.

Let M be a von Neumann algebra determined in the above manner by a freely-acting group G of MP automorphisms of a hyperstonian measure space. Then if G is abelian (or more generally, if finitely generated subgroups of G satisfy the conditions of Theorem 1), the algebra M will be approximately finite, this by Corollary 6.1. This generalizes a well-known result of Murray and von Neumann ([4, Lemma 5.2.3, stated without proof]).

University of California,
Los Angeles, California.

REFERENCES.

[1] J. Dixmier, *Les algèbres d'opérateurs dans l'espace Hilbertien*, Paris, 1957.

[2] ———, "Sous-anneaux abéliens maximaux dan les facteurs de type fini," *Annals of Mathematics*, vol. 59 (1954), pp. 279-286.

[3] H. A. Dye, "On groups of measure preserving transformations. I," *American Journal of Mathematics*, vol. 81 (1959), pp. 119-159.

[4] F. J. Murray and J. von Neumann, "On rings of operators. IV," *Annals of Mathematics*, vol. 44 (1943), pp. 716-808.

[5] H. Umegaki, "Conditional expectation in an operator algebra," *Tôhoku Journal of Mathematics*, vol. (2) 6 (1954), pp. 177-181.

[6] H. Widom, "Approximately finite algebras," *Transactions of the American Mathematical Society*, vol. 83 (1956), pp. 170-178.

ON HOMOGENEOUS AFFINE SPACES OF LINEAR ALGEBRAIC GROUPS.*

By A. Białynicki-Birula.[1]

0. Let G be a linear algebraic group and let H be a closed subgroup of G. In this paper we give some necessary and sufficient conditions for G/H to be affine. For example, one of our results (Theorem 1) says that if the base field is of characteristic 0 and G is fully reducible then G/H is affine if and only if H is fully reducible. In the case where the base field is the field of complex numbers the theorem is known (see Theorem 3.5 [3]) but the proof presented here is based on a different idea.

The author would like to express his thanks to G. P. Hochschild for his advice in the preparation of this paper.

1. Let the ground field K be algebraically closed. All algebraic sets, algebraic groups and morphisms that we consider are supposed to be defined over k. Let G be a connected linear algebraic group. Then $k(G)$ denotes the field of all rational functions of G (defined over k) and $k[G]$ denotes the ring of all regular (i.e., representative) functions of G (defined over k). If $f \in k(G)$ and $g \in G$ the left and right translates $g \cdot f$, $f \cdot g$ are defined by

$$(g \cdot f)(x) = f(xg), \qquad (f \cdot g)(x) = f(gx).$$

For any subgroup H of G, H_l, H_r denote the groups of all left and right translations of $k(G)$ by elements of H.

If H is a closed subgroup of G then G/H denotes the algebraic variety of right H-cosets of G. The field of rational functions $k(G/H)$ of G/H is canonically isomorphic to the subfield $k(G)^{H_l}$ composed of all functions from $k(G)$ that are fixed under H_l. Moreover, the ring $k[G/H]$ of regular functions of G/H is canonically isomorphic to $k[G]^{H_l} = k[G] \cap k(G)^{H_l}$. We shall identify $k(G/H)$ with $k(G)^{H_l}$ and $k[G/H]$ with $k[G]^{H_l}$ according to these canonical isomorphisms. The subgroup H is called *observable* (in G) if $k[G]^{H_l}$ generates $k(G)^{H_l}$ as a field [2]. It is known that H is observable if and only if G/H is quasi-affine [2]. The field $k(G)^{H_l}$ is stable under G_r; this action of G on $k(G)^{H_l}$ is induced by the transitive action of G on G/H.

* Received March 6, 1963.

[1] Written while the author was partially supported by N. S. F. Grant G-24943.

If $f \in k[G]^{H_l}$ then $(G/H)_f$ denotes the set of all points $x \in G/H$ such that $f(x) \neq 0$.

Suppose that G acts as a group of linear automorphisms on some finite dimensional vector space M over k. We shall say that M is a rational finite dimensional G-module if the representation of G in M is rational. M^G denotes the subspace of M composed of all elements $v \in M$ such that $g(v) = v$, for every $g \in G$. One can prove (see Lemma 2 in [9]) that if G acts on an affine space V then there exists a G-embedding of V onto a closed subset of a rational finite dimensional G-module.

2. We shall say that a subring R of $k[G]$ is G_r-simple if it contains k, is stable under G_r and has no G_r-stable ideal other than (0) and R.

PROPOSITION 1. *Let H be an observable subgroup of G. Then G/H is affine if and only if $k[G]^{H_l}$ is G_r-simple.*

Proof. If G/H is affine then $k[G]^{H_l}$ is G_r-simple since G/H is a homogeneous G-space.

Now, assume that $k[G]^{H_l}$ is G_r-simple. We may find a non-zero function $f \in k[G]^{H_l}$ such that $(G/H)_f$ is affine (since G/H is quasi-affine). Since $k[G]^{H_l}$ is G_r-simple, there exist an integer n, some elements $g_1, \cdots, g_n \in G$ and $h_1, \cdots, h_n \in k[G]^{H_l}$ such that $\sum_i (f \cdot g_i) h_i = 1$. We have $(G/H)_{f \cdot g_i} = g_i^{-1}(G/H)_f$. Therefore $(G/H)_{f \cdot g_1}, \cdots, (G/H)_{f \cdot g_n}$ form an open and affine covering of G/H. Thus G/H is affine (see, e.g., Theorem on p. 05 in [4]).

We can also show that every G_r-simple subring R of $k[G]$, such that $k(G)$ is separable over the field generated by R, is of the form $k[G]^{H_l}$, for some closed subgroup H of G. Indeed, suppose that R satisfies the above conditions. Then the field generated by R is G_r-stable and is of the form $k(G)^{H_l}$, for some closed subgroup H of G (see [1]). Now, $k[G]^{H_l}$ contains R and is G_r-simple (since every $f \in k[G]^{H_l}$ can be written as f_1/f_2, where $f_1, f_2 \in R$). By Proposition 1, G/H is therefore affine, and hence $k[G]^{H_l}$ is finitely generated. Thus we can find a non-zero function $h \in R$ such that every element $f \in k[G]^{H_l}$ can be written as f_l/h^m, where $f \in R$ and m is a non-negative integer. Next, we can find an integer n and elements $g_1, \cdots, g_n \in G$ and $q_1, \cdots, q_n \in R$ such that $\sum (h \cdot g_i) q_i = 1$, since $h \neq 0$ and R is G_r-simple. In order to prove that $R = k[G]^{H_l}$, it remains to show that $k[G]^{H_l} \subset R$. Let $f \in k[G]^{H_l}$. Then we can choose an integer r and elements $h_1, \cdots, h_n \in R$ such that

$$f \cdot g_i^{-1} = \frac{h_i}{h^r}. \text{ This gives } f = \frac{h_i \cdot g_i}{(h \cdot g_i)^r}, \text{ for } i = 1, \cdots, n.$$

But, since $\sum(h\cdot g_i)q_i = 1$, we have $(\sum(h\cdot g_i)q_i)^{p^r} = 1$ and we can find $q_1^*,\cdots,q_n^* \in R$ such that $\sum(h\cdot g_i)^r q_i^* = 1$. Hence

$$f = f\cdot\sum(h\cdot g_i)^r q_i^* = \sum f(h\cdot g_i)^r q_i^* = \sum(h_i\cdot g_i)q_i^*.$$

Thus $f \in R$.

The following proposition follows from the above remark and Proposition 1.

PROPOSITION 2. *The mapping $H \to k[G]^{H_l}$ defines a one-one Galois correspondence between the family of all closed subgroups H of G such that G/H is affine and the family of all G_r-simple subrings R of $k[G]$ such that $k(G)$ is separable over the field generated by R. In particular, if k is of characteristic 0, this is a correspondence between the family of all closed subgroups H of G such that G/H is affine and the family of all G_r-simple subrings of $k[G]$.*

3. We assume now that k is of characteristic 0. It is known (see Theorem 5.1 in [7]) that

(i) If H is a fully reducible closed subgroup of G then G/H is affine.

This result can be proved in the following way. If H is a fully reducible subgroup of G then there exists a $k[G]^{H_l}$-linear projection $\phi\colon k[G] \to k[G]^{H_l}$. Hence $k[G]^{H_l}$ is G_r-simple. Indeed, if $f \neq 0$, $f \in k[G]^{H_l}$ then there exist an integer n, elements $g_1,\cdots,g_n \in G$ and $f_1,\cdots,f_n \in k[G]$ such that $\sum(f\cdot g_i)f_i = 1$, since $k[G]$ is G_r-simple. Hence

$$\sum(f\cdot g_i)\phi(f_i) = \phi(\sum(f\cdot g_i)\cdot f) = 1.$$

Moreover, one can show (see Proposition 2.1 in [7]) that H is observable. Thus it follows from Proposition 1 that G/H is affine.

The theorem (i) is a particular case of the following result of D. Mumford (see his forthcoming book *Geometric Invariant Theory*).

(ii) Let X be an affine algebraic variety and let G be a reductive linear algebraic group. Suppose that G acts on V and assume that all orbits are closed. Then the orbit variety X/G exists and is affine.

In the sequel we shall also use the following lemma.

LEMMA 1. *Let G be a linear algebraic group and let H_1, H_2 be two closed subgroups of G such that $H_1 \supset H_2$. If H_1/H_2 is affine then the canonical morphism $\pi\colon G/H_2 \to G/H_1$ is affine, i.e., there exists a covering $\{U_\alpha\}$ of G/H_1 by affine open sets such that $\pi^{-1}(U_\alpha)$ is affine, for any α.*

Proof. It suffices to find an open affine and non-empty subset U of G/H_1 such that $\pi^{-1}(U)$ is affine, since π commutes with the action of G on G/H_2 and G/H_1. Denote by π_1, π_2 the canonical morphisms of G onto G/H_1, G/H_2, respectively. Then $\pi\pi_2 = \pi_1$. Moreover, $\pi_1 \colon G \to G/H_1$ is a principal H_1-fibre space (Proposition 3 in [9]). Hence one can find an open affine subset U of G/H_1 such that $U \neq \emptyset$ and $\pi_1 \mid \pi_1^{-1}(U) : \pi_1^{-1}(U) \to U$ is isotrivial, i. e., there exists an affine variety U' and an integral morphism $p \colon U' \to U$ such that the induced H_1-fibre space $U' \times_U \pi_1^{-1}(U) \to U'$ is trivial (i. e. equivalent to the fibre space $U' \times H_1 \to U'$ with the natural projection onto U' and with the action of H_1 defined by $h_1(u,h) = (u, hh_1)$, for any $(u,h) \in U' \times H_1$, $h_1 \in H_1$). Now, we have the following commutative diagram

$$\begin{tikzcd}
U' \times_U \pi_1^{-1}(U) \arrow[r, "p_1"] \arrow[d, "\pi_2^*"] & \pi_1^{-1}(U) \arrow[d, "\pi_2 \mid \pi_1^{-1}(U)"] \\
U' \times_U \pi_1^{-1}(U) \arrow[r, "p'"'] \arrow[d, "\pi^*"] & \pi^{-1}(U) \arrow[d, "\pi \mid \pi^{-1}(U)"] \\
U' \arrow[r, "p"'] & U
\end{tikzcd}$$

with p', p_1, π^*, π_2^* induced by p, $\pi \mid \pi^{-1}(U)$, $\pi_2 \mid \pi_1^{-1}(U)$. p' is integral since p is integral and p' is induced by p and change of the base variety $\pi \mid \pi^{-1}(U) : \pi^{-1}(U) \to U$. Hence it follows from a theorem of Chevalley (Theorem on p. 136 in [6]) that if $U' \times_U \pi^{-1}(U)$ is affine then $\pi^{-1}(U)$ is also affine. But $U' \times_U \pi^{-1}(U)$ is the H_2-orbit space of $U' \times \pi_1^{-1}(U) \approx U' \times H_1$ and hence is isomorphic to $U' \times H_1/H_2$ and thus is affine since U' and H_1/H_2 are affine.

COROLLARY 1. *Let G, H_1, H_2 be as in the lemma. If H_1/H_2 is affine and G/H_1 is affine then also G/H_2 is affine.*

Proof. It suffices to notice that if π is the canonical morphism $G/H_2 \to G/H_1$ then, for any affine open subset $U \subset G/H_1$, $\pi^{-1}(U)$ is affine, since π is affine.

Let G_u, H_u denote the maximal unipotent normal subgroups of G and H, respectively. (see [8])

PROPOSITION 3. *Let G be a linear algebraic group and let H be a closed subgroup of G. Then G/H is affine if and only if G/H_u is affine.*

Proof. Suppose that G/H is affine then it follows from Corollary 1 that G/H_u is also affine. Now, assume that G/H_u is affine. The fully reducible group H/H_u acts on G/H_u from the right side and G/H is the orbit space. Hence it follows from (ii) that G/H is affine.

COROLLARY 2. *If $H_u \subset G_u$ then G/H is affine.*

Proof. If $H_u \subset G_u$ then G_u/H_u is affine (see e.g., Proposition 2 in [5]). Hence, by Corollary 1, G/H_u is affine. Thus it follows from Proposition 3 that G/H is affine.

We do not know any group-theoretic necessary and sufficient condition on G and H for G/H to be affine. One can easily see that the condition given in Corollary 2 is not necessary. But such a condition can be found if G is fully reducible. Namely, we have the following result.

THEOREM 1. *Let G be a fully reducible linear algebraic group and let H be a closed subgroup of G. Then the following conditions are equivalent*

(a) *G/H is affine*

(b) *H is fully reducible*

(c) *any finite-dimensional rational H-module M is an H-submodule of a finite-dimensional rational G-module N such that $M^H = N^G$.*

Proof. We have sketched above a proof (b) $\Rightarrow$ (a). (c) $\Rightarrow$ (b). Suppose (c). In order to prove that H is fully reducible it suffices to show that, for any finite dimensional rational H-module M, M^H is a direct summand of M. But, this follows from (c) and the fact that G is fully reducible.

(b) $\Rightarrow$ (c). If H is fully reducible and M is a finite-dimensional rational H-module then M^H is a direct summand of M. Let $M = M^H + M_1$. Since H is fully reducible it is observable and there exists a finite dimensional rational G-module N_1 such that M_1 is an H-submodule of N_1 ([2]). We may assume that $N_1{}^G = (0)$. Now, consider $N = M^H + N_1$ as a G-module with the trivial action of G on M^H. Then N satisfies conditions of (c).

(a) $\Rightarrow$ (b). In the proof we shall use the following known results

(iii) (Theorem 10 in [9], cf. also Proposition 14 in [9]). If H_0 is a closed connected and unipotent subgroup of G then the canonical morphism $\pi_0\colon G \to G/H_0$ defines a locally trivial principal fibre space with the group H_0.

(iv) (Proposition 1 in [5]). If $p_0\colon V \to V_0$ is a locally trivial principal fibre space with the group H_0, where H_0 is as in (iii) and V_0 is affine then the fibering is trivial.

Suppose (a). It follows from Proposition 3 that G/H_* is affine and it suffices to prove that H_* is fully reducible. Hence we may assume that $H = H_*$. By (iii) and (iv) we obtain that G is isomorphic to $(G/H) \times H$ under an isomorphism commuting with right multiplications by elements of H (if $(a, h) \in (G/H) \times H$ and $h_1 \in H$ then we define $(a, h)h_1 = (a, hh_1)$). Hence the isomorphism induces an injective homomorphism i of $k[H]$ into $k[G]$. i sends constant functions onto constant functions and commutes with elements of H. Since every finite-dimensional rational H-module is isomorphic to a submodule of a direct sum of H-modules $k[H]$, where H acts by left translations, hence H satisfies (c). Thus H is fully reducible.

UNIVERSITY OF CALIFORNIA,
INSTITUTE OF MATHEMATICS P. A. N., WARSAW, POLAND.

REFERENCES.

[1] E. Abe and T. Kanno, "Some remarks on algebraic groups," *Tôhoku Mathematical Journal*, vol. 11 (1959), pp. 376-384.

[2] A. Bialynicki-Birula, G. Hochschild and G. Mostow, "Extensions of representations of linear algebraic groups," *American Journal of Mathematics*, vol. 85 (1963), pp. 131-144.

[3] A. Borel and Harish-Chandra, "Arithmetic subgroups of algebraic groups," *Annals of Mathematics*, vol. 75 (1962), pp. 485-535.

[4] P. Gabriel, "Le théorèm de Serre," *Séminaire C. Chevalley*, t. 3 (1958/59), (exposé no. 2).

[5] A. Grothendieck, "Torsion homologique et section rationnelles," *Séminaire C. Chevalley*, t. 2, *Anneaux de Chow et application* (1958) (exposé no. 1).

[6] ———, "Éléments de géométrie algébrique, 2," *Institut des Hautes Études Scientifiques, Publications Mathématiques* no. 8, Paris, 1961.

[7] G. Hochschild and B. Kostant, "Differential forms and Lie algebra cohomology for algebraic linear groups," *Illinois Journal of Mathematics*, vol. 6 (1962), pp. 264-281.

[8] G. Mostow, "Fully reducible subgroups of algebraic groups," *American Journal of Mathematics*, vol. 78 (1956), pp. 200-221.

[9] M. Rosenlicht, "Some basic theorems on algebraic groups," *American Journal of Mathematics*, vol. 78 (1956), pp. 401-443.

[10] J-P. Serre, "Espaces fibrés algébriques," *Séminaire C. Chevalley*, t. 2, *Anneaux de Chow et application* (1958) (exposé no. 1).

ON PRIME J-RINGS WITH UNIFORM ONE-SIDED IDEALS.*

By YUZO UTUMI.

0. We denote by $l(A, B)$ the set of elements $x \in A$ such that $xB = 0$. Similarly $r(A, B)$ denotes the set of $x \in A$ with $Bx = 0$.

A ring S is called a J-ring if it satisfies the following two conditions:

(J_l) If $r(S, A) \neq 0$ for a left ideal A of S, there exists a nonzero left ideal B of S such that $A \cap B = 0$.

(J_r) = the right-left symmetry of (J_l).

We say that a ring S is a J_l-ring if it fulfills (J_l). A J_r-ring is defined in an obvious way.

A ring S is called a left quotient ring (in the sense of R. E. Johnson) of a subring R if every nonzero left R-submodule of S has a nonzero intersection with R. A right quotient ring is defined similarly. An extension ring S of a ring R is called a two-sided quotient ring of R if S is a left quotient ring of R, and also a right quotient ring of R.

For any J_l-ring S there exists such a left quotient ring $\bar{S}_l$ that every left quotient ring of S has an isomorphic, over S, image in $\bar{S}_l$. $\bar{S}_l$ is unique up to isomorphism over S, and is called the maximal left quotient ring of S. Similarly the maximal right quotient ring $\bar{S}_r$ exists for any J_r-ring S. In case S is a J-ring, it may be shown that S has the maximal two-sided quotient ring $\bar{S}_t$.

A ring S is called a continuous transformation ring if there is a pair of dual vector spaces (V, V') such that S is the ring of continuous linear transformations of the vector space V topologized by the V'-topology. This concept is right-left symmetrical; that is, S may be regarded as the ring of continuous linear transformations of the vector space V' topologized by the V-*topology*. (*See* [2; §§ 6, 7, Chap. IV]).

A left ideal A of a ring S is called uniform if we have $B \cap C \neq 0$ for any nonzero left ideals B and C such that $B, C \subset A$. A uniform right ideal is defined similarly.

The main theorem of this paper is the following: Let S be a prime J-ring

* Received March 13, 1963.

with a uniform left ideal and a uniform right ideal. Then $\bar{S}_l$ is a continuous transformation ring. Suppose $\bar{S}_l$ is determined by a pair of dual vector spaces (V, V'). $\bar{S}_l$ ($\bar{S}_r$ resp.) is the ring of linear transformations of V (V' resp.).

1. LEMMA 1.1. *Let S be a ring, and Se a minimal left ideal generated by an idempotent e. Let A be a right Ore domain, and suppose that $Se \supset A$. Then there is an idempotent f such that $Se = Sf$ and $fSf \supset A$.*

Proof. Let $0 \neq a \in A$. Then $a^2 \neq 0$, and $ea \neq 0$. Hence $Sea = Se$, and so $bea = a$ for some b. Set $f = e + be - ebe$. Then $f = f^2$, $Se = Sf$, and $a \in fSf$. Let $0 \neq x \in A$. Since A is a right Ore domain, there are a', $a'' \in A$ such that $xa' = aa'' \neq 0$. Now fSf is the endomorphism ring of the S-module $Sf = Se$, and hence it is a division ring. Since $a' \neq 0$, $a'^2 \neq 0$, and $fa' \neq 0$. Thus we may find an element c such that $(fa')c = f$. $x = xf = xfa'c = xa'c = aa''c \in fS$. Therefore $x = fxf$, which shows that $A \subset fSf$, as desired.

A left (resp. right) ideal A of a ring S is called large if A has a nonzero intersection with every nonzero left (resp. right) ideal of S.

If a ring S is a left (resp. right) quotient ring of a ring R, then it is known that for any element x of S there exists a large left (resp. right) ideal A of R such that $Ax \subset R$ (resp. $xA \subset R$).

It is easy to see that the conditions (J_l) and (J_r) are equivalent to the following (J'_l) and (J'_r) respectively.

(J'_l) $r(S, A) = 0$ for any large left ideal A of S.
(J'_r) = the right-left symmetry of (J'_l).

LEMMA 1.2. *Let R be a J_l-ring, and A a large left ideal of R. Let e be an idempotent in a left quotient ring S of R, and let $x \in R$. If $x = ex$ and $(Se \cap A)x = 0$, then $x = 0$.*

Proof. It is known and easily seen that $(B =)$ $(Se \cap A) + (S(1-e) \cap A)$ is a large left ideal of R. By assumption $Bx = 0$. Since R is a J_l-*ring*, it follows that $x = 0$, as desired.

LEMMA 1.3. *Let S be a two-sided quotient ring of a semiprime J-ring R, and let e be an idempotent of S. Then eSe is a two-sided quotient ring of $eSe \cap R$. $eSe \cap R$ is semiprime. In case R is prime, $eSe \cap R$ is also prime.*

Proof. Let $0 \neq x \in eSe$. Then $Ax \subset R$ for some large left ideal A of R. $(eS \cap R)(Se \cap A) \subset eSe \cap R$, and moreover

$$(eS \cap R)(Se \cap A)x \subset eSex \cap RAx \subset eSe \cap R.$$

Since $x \neq 0$, $(Se \cap A)x \neq 0$ by Lemma 1.2, and so $(Se \cap A)x(eS \cap R) \neq 0$

by the right-left symmetry of Lemma 1.2, and hence $((Se \cap A)x(eS \cap R))^2 \neq 0$, which implies that $(eS \cap R)(Se \cap A)x \neq 0$. Therefore eSe is a left quotient ring of $eSe \cap R$. Similarly eSe is a right quotient ring of $eSe \cap R$.

Next, we suppose that $BC = 0$ for two ideals B and C of $eSe \cap R$. Then

$$0 = (eSe \cap R)B(eSe \cap R)C \supset (eS \cap R)(Se \cap R)B(eSe \cap R)C.$$

Hence $(Se \cap R)B(eSe \cap R)C(eS \cap R) = 0$ by the semiprimeness of R, and so $(Se \cap R)B(eSe \cap R)C = 0$ by the symmetry of Lemma 1.2, whence $(Se \cap R)B(eS \cap R)(Se \cap R)C = 0$. By the semiprimeness

$$(Se \cap R)C(Se \cap R)B(eS \cap R) = 0,$$

and hence $(Se \cap R)C(Se \cap R)B = 0$ by the symmetry of Lemma 1.2.

Case 1. Let $B = C$. Then, since R is semiprime, we have $(Se \cap R)B = 0$, and therefore $B = 0$ by Lemma 1.2.

Case 2. Suppose R is prime. Then we have $(Se \cap R)B = 0$ or $(Se \cap R)C = 0$, and hence $B = 0$ or $C = 0$ by Lemma 1.2, completing the proof.

LEMMA 1.4. *Let S be a J-ring. Then S has the maximal two-sided quotient ring S_t. S_t may be regarded as the subring of $\bar{S}_l$ consisting of elements x such that the set of $y \in S$ with $xy \in S$ forms a large right ideal of S.*

Proof. By [6; Lemma 3.1] the set A of $x \in \bar{S}_l$ such that $\{y \in S : xy \in S\}$ is large is a right quotient ring of S. Since $\bar{S}_l \supset A \supset S$, A is a left quotient ring of S, too. Let B be any two-sided quotient ring of S. Then B may be looked upon as a ring between $\bar{S}_l$ and S. Since $\{y \in S : xy \in S\}$ is a large right ideal of S for every $x \in B$, we have $B \subset A$, as desired.

Remark. Right-left symmetrically we may find a subring C of $\bar{S}_r$, which is the maximal two-sided quotient ring of S. In view of the following lemma C may be identified with the subring of $\bar{S}_l$ in Lemma 1.4.

LEMMA 1.5. *Let S be a J-ring, and suppose that both P and Q are the maximal two-sided quotient rings of S, then there is an isomorphism, over S, of P and Q.*

Proof. There exists an isomorphism, over S, of P onto a subring of Q. Also we may find an isomorphism, over S, of Q onto a subring of P. The product of these isomorphisms is an isomorphism of P onto a subring of P. Thus, the lemma follows from the following

LEMMA 1.6. *Let S be a J_l-ring, and P a left quotient ring of S. Then any homomorphism, over S, of P into itself is the identity mapping.*

Proof. Let v be a homomorphism, over S, of P into a subring of P, and let $x \in P$. Then $Ax \subset S$ for some large left ideal A of S. Let $a \in A$. $a(v(x)) = (v(a))(v(x)) = v(ax) = ax$. Hence $A(v(x) - x) = 0$. Denote by B the singular submodule of the left S-module P, that is, the set of $y \in P$ such that $l(S, y)$ is a large left ideal. Then $v(x) - x \in B$. Since B is a left S-module and $B \cap S = 0$, we may conclude that $B = 0$. Thus, $x = v(x)$, as desired.

2. In this section we always suppose that a ring T is a prime J-ring with a uniform left ideal and a uniform right ideal.

A ring S is called a right full linear ring if it is the ring of linear transformations of a vector space V over a division ring, V being regarded as a right S-module. Similarly, a left full linear ring is defined.

It is evident from the definitions that $\bar{S}_l$ is prime for any prime J_l ring S. Thus, in particular $\bar{T}_l$ is a prime ring. From the assumption that T contains a uniform left ideal it follows that $\bar{T}_l$ has a minimal left ideal. Thus, $\bar{T}_l$ is a primitive ring with nonzero socle, and hence it is a right full linear ring by [5; (5.1)]. Similarly, $\bar{T}_r$ is a left full linear ring. Let $\bar{T}_t$ be the maximal two-sided quotient ring of T. We may suppose that $\bar{T}_l \supset \bar{T}_t$ and $\bar{T}_r \supset \bar{T}_t$ by Lemma 1.4. In this section we shall show that $\bar{T}_t$ is a primitive ring with nonzero socle.

Let e be a primitive idempotent of $\bar{T}_l$. Though the following Lemmas 2.1, 2.2 and 2.3 are essentially the same with [4; 2.1, 2.2], we shall show their proofs for the sake of completeness.

LEMMA 2.1. *There exists a primitive idempotent f of $\bar{T}_r$ such that $(\bar{T}_l e \cap T)(f\bar{T}_r \cap T) \neq 0$ and $(\bar{T}_l e \cap T) \cap (f\bar{T}_r \cap T) \neq 0$.*

Proof. If $(\bar{T}_l e \cap T)(f'\bar{T}_r \cap T) = 0$ for every primitive idempotent f' of $\bar{T}_r$, then $(\bar{T}_l e \cap T)f'\bar{T}_r = 0$ in $\bar{T}_r$. Thus, $(\bar{T}_l e \cap T)P = 0$ where P denotes the socle of $\bar{T}_r$. Therefore, $(\bar{T}_l e \cap T)(P \cap T) = 0$, contradicting the primeness of T. This shows that $(\bar{T}_l e \cap T)(f\bar{T}_r \cap T) \neq 0$ for some primitive idempotent f of $\bar{T}_r$. By the semiprimeness of T we have $(f\bar{T}_r \cap T)(\bar{T}_l e \cap T) \neq 0$, whence $(f\bar{T}_r \cap T) \cap (\bar{T}_l e \cap T) \neq 0$, as desired.

LEMMA 2.2. *Let $0 \neq x \in \bar{T}_l e \cap T$ and $0 \neq y \in f\bar{T}_r \cap T$. Then $xy \neq 0$.*

Proof. If $xy = 0$, then $r(\bar{T}_r, x) \cap f\bar{T}_r \neq 0$. Hence $r(\bar{T}_r, x) \supset f\bar{T}_r$, and so $x(f\bar{T}_r \cap T) = 0$. It follows from this that $l(\bar{T}_l, f\bar{T}_r \cap T) \cap \bar{T}_l e \neq 0$, and so $l(\bar{T}_l, f\bar{T}_r \cap T) \supset \bar{T}_l e$, whence $(\bar{T}_l e \cap T)(f\bar{T}_r \cap T) = 0$, contradicting Lemma 2.1.

LEMMA 2.3. $(\bar{T}_l e \cap T) \cap (f\bar{T}_r \cap T) (= A)$ *is an Ore domain.*

Proof. By Lemma 2.2 A has no zero divisors. Let $0 \neq p \in A$ and $0 \neq q \in A$. It follows from the general theory of quotient rings that $\bar{T}_l e \cap T$ is a uniform left ideal of T. Hence $(\bar{T}_l e \cap T)p \cap (\bar{T}_l e \cap T)q \neq 0$, and so $((\bar{T}_l e \cap T)p \cap (\bar{T}_l e \cap T)q)(f\bar{T}_r \cap T) \neq 0$ by Lemma 2.2. Since T is semi-prime, $(f\bar{T}_r \cap T)\cdot((\bar{T}_l e \cap T)p \cap (\bar{T}_l e \cap T)q) \neq 0$. Thus,

$$0 \neq (f\bar{T}_r \cap T)(\bar{T}_l e \cap T)p \cap (f\bar{T}_r \cap T)(\bar{T}_l e \cap T)q \subset Ap \cap Aq,$$

and hence $Ap \cap Aq \neq 0$. Similarly, $pA \cap qA \neq 0$. This shows that A is an Ore domain, as desired.

LEMMA 2.4. *There exists an idempotent g in $\bar{T}_t$ with the properties that* (i) *$g\bar{T}_t g$ is the quotient division ring of A,* (ii) *$\bar{T}_l e = \bar{T}_l g$, and* (iii) *$f\bar{T}_r = g\bar{T}_r$.*

Proof. Since A is an Ore domain, and $A \subset \bar{T}_l e$, there is an idempotent g of $\bar{T}_l$ such that $\bar{T}_l g = \bar{T}_l e$, $g\bar{T}_l g \supset A$ by Lemma 1.1. Let K be the quotient ring of the Ore domain A. Then, since $g\bar{T}_l g$ is a division ring, we may suppose that $g\bar{T}_l g \supset K \supset A$. Thus the identity element of K is g, and hence $g \in K$.

On the other hand, there exists an idempotent h of $\bar{T}_r$ such that $h\bar{T}_r = f\bar{T}_r$ and $h\bar{T}_r h \supset A$ by the right-left symmetry of Lemma 1.1.

$\bar{T}_t$ may be regarded as a subring of $\bar{T}_l$ by Lemma 1.4. We shall show that K is contained in the subring $\bar{T}_t$ of $\bar{T}_l$. To see this it is enough to prove that for any $0 \neq t \in K$ there is a large right ideal C of T such that $tC \subset T$. Now there exists an element $a \in A$ with $0 \neq ta \in A$. Hence $ta(hT_r \cap T) \subset T$. Let $0 \neq b \in A$. Then $b'b = g$ for some $b' \in K \subset \bar{T}_l$. We have

$$t((1-h)\bar{T}_r \cap T) = tb'b((1-h)\bar{T}_r \cap T)$$

in $\bar{T}_l$. However $b((1-h)\bar{T}_r \cap T) \subset h\bar{T}_r h(1-h)\bar{T}_r = 0$ in $\bar{T}_r$. Therefore $t((1-h)\bar{T}_r \cap T) = 0$ in $\bar{T}_l$. Set $C = a(h\bar{T}_r \cap T) + ((1-h)\bar{T}_r \cap T)$. Then C is a right ideal of T, and $tC \subset T$. Now $(h\bar{T}_r \cap T) \cap ((1-h)\bar{T}_r \cap T)$ is a large right ideal of T, and $h\bar{T}_r \cap T$ is a uniform right ideal of T. $a(h\bar{T}_r \cap T) \subset A(h\bar{T}_r \cap T) \subset (f\bar{T}_r \cap T)^2 \subset f\bar{T}_r \cap T = h\bar{T}_r \cap T$. And moreover $a(h\bar{T}_r \cap T) = a(f\bar{T}_r \cap T) \neq 0$ by Lemma 2.2. Thus, it follows that C is also a large right ideal of T, as desired. Therefore $K \subset \bar{T}_t \subset \bar{T}_l$.

Next, we regard $\bar{T}_t$ as a subring of $\bar{T}_r$: $K \subset \bar{T}_t \subset \bar{T}_r$. Let b be any nonzero element of A. Then $g = bb'$ for some $b' \in K$, since K is the quotient division ring of A. $0 \neq g\bar{T}_r = bb'\bar{T}_r \subset b\bar{T}_r \subset (f\bar{T}_r \cap T)\bar{T}_r \subset f\bar{T}_r$. By the minimality of $f\bar{T}_r$ we have $g\bar{T}_r = f\bar{T}_r$.

$$K \cap T \subset g\bar{T}_t g \cap T \subset (\bar{T}_t g \cap T) \cap (g\bar{T}_t \cap T) \subset (\bar{T}_l g \cap T) \cap (g\bar{T}_r \cap T)$$
$$= (\bar{T}_l e \cap T) \cap (f\bar{T}_r \cap T) = A \subset K \cap T.$$

Therefore $A = g\bar{T}_t g \cap T$.

Now we have $A \subset K \subset g\bar{T}_t g$. By Lemma 1.3 $g\bar{T}_t g$ is a two-sided quotient ring of $g\bar{T}_t g \cap T = A$, and therefore $g\bar{T}_t g$ is a two-sided quotient ring of K. However K is a division ring, and has no proper quotient rings. Thus, $K = g\bar{T}_t g$, completing the proof.

LEMMA 2.5. *$\bar{T}_t$ is a primitive ring with nonzero socle.*

Proof. From the primeness of T it follows immediately that $\bar{T}_t$ is also a prime ring. By [2; Proposition 1, p. 65] $\bar{T}_t g$ is a minimal left ideal of $\bar{T}_t$ since $g\bar{T}_t g = K$ is a division ring. Therefore $\bar{T}_t$ is a primitive ring with nonzero socle, as desired.

3. Let P be a primitive ring with nonzero socle, and let d be a primitive idempotent of P. Then, as is well known, dP is a vector space over dPd, and P is a dense ring of linear transformations of the vector space dP. Similarly, P is a dense ring of linear transformations of the vector space Pd over dPd. The correspondence $(x, y) \to xy$ gives a bilinear transformation $(dP, Pd) \to dPd$. (dP, Pd) is a pair of dual vector spaces with respect to this bilinear transformation. dP is a topological vector space over dPd by the Pd-topology. A subspace of dP is closed if and only if it is the orthogonal complement of a subspace of Pd. Similarly Pd is also a topological vector space by the dP-topology. Let X be the ring of continuous linear transformations of Pd. Then X may be regarded as the ring of continuous linear transformations of Pd. X contains P, and the socle of X coincides with that of P. (See [2; p. 77]).

Let Y be the ring of linear transformations of dP. Then $Y \supset X \supset P$. Similarly, the ring Z of linear transformations of Pd contains X.

In view of [5; (5.1)] $Y = \bar{P}_l$ and $Z = \bar{P}_r$. Moreover, $X = \bar{P}_t$ by [5; (5.3)].

THEOREM 3.1. *Let T be a prime J-ring with a uniform left ideal and a uniform right ideal. Then $\bar{T}_t$ is a continuous transformation ring.*

Proof. By virtue of Lemma 2.5 $\bar{T}_t$ is a primitive ring with nonzero socle. Set $P = \bar{T}_t$. Then it is evident from the general theory of quotient rings that $\bar{P}_t = \bar{T}_t$. Since $X = \bar{P}_t$, $\bar{T}_t = X$ is a continuous transformation ring, as desired.

Remark. $Y = \bar{P}_l = \bar{T}_l$ and $Z = \bar{P}_r = \bar{T}_r$.

A left ideal A of a ring S is called closed if for any left ideal A' with $A \subset A'$, $A \neq A'$ there exists a nonzero left ideal A'' of S such that $A'' \subset A'$ and $A'' \cap A = 0$. (Closed (R. E. Johnson) = complemented (Utumi [5]) = complement (Goldie [1])). It is known that the set of closed left ideals of a J_l-ring S forms a complete complemented modular lattice $L(S)$. If S' is a left quotient ring of a J_l-ring S, then S' is also a J_l-ring, and the correspondence $A \in L(S') \rightarrow A \cap S$ gives an isomorphism $L(S') \rightarrow L(S)$. A left ideal of $\bar{S}_l$ for a J_l-ring S is closed if and only if it is generated by an idempotent.

THEOREM 3.2. *Let T be a prime J-ring with uniform left ideal and a uniform right ideal. Let e be an idempotent in the socle of $\bar{T}_t$. Then $e\bar{T}_te$ is a simple ring with minimum condition, and is the classical quotient ring of $e\bar{T}_te \cap T$.*

Proof. It is evident that $e\bar{T}_te$ is the total matrix ring of finite degree over a division ring. By Lemma 1.3 $e\bar{T}_te \cap T$ is a prime ring, and $e\bar{T}_te$ is a two-sided quotient ring of $e\bar{T}_te \cap T$. Since $e\bar{T}_te$ is a J-ring, it is not difficult to see that $e\bar{T}_te \cap T$ is also a J-ring. In view of the isomorphisms between the lattices of closed one-sided ideals of $e\bar{T}_te$ and $e\bar{T}_te \cap T$ it follows that $e\bar{T}_te \cap T$ satisfies the maximum conditions for closed one-sided ideals. Moreover, since any annihilator one-sided ideal of the J-ring $e\bar{T}_te \cap T$ is closed, $e\bar{T}_te \cap T$ satisfies the maximum conditions for annihilator one-sided ideals too. Therefore $e\bar{T}_te \cap T$ has the classical quotient ring Q by [1; Theorem 11],and Q is a simple ring with minimum condition by [1; Theorem 12]. Therefore $e\bar{T}_te$ may be identified with Q since both Q and $e\bar{T}_te$ are the maximal (two-sided) quotient ring of $e\bar{T}_te$, completing the proof.

The following proposition may be interesting because of the importance of the isomorphism between the lattices of closed one-sided ideals.

PROPOSITION 3.3. *Let A be a left ideal of the continuous transformation ring X. Then the following conditions are equivalent:*

(i) *A is a finite dimensional element of $L(X)$.*

(ii) *A is generated by an idempotent in the socle of X.*

(iii) *A is a left ideal of Z generated by an idempotent in the socle of X.*

(iv) *$A = Ye \cap X$ where e is an idempotent in the socle of Y.*

Proof may be easily obtained from the theory of primitive rings, and will be omitted.

4. LEMMA 4.1. *Let S be a ring. Suppose that there is given a family F of subsets of S with the following properties:*

(1) *Every nonzero annihilator right ideal contains a member of F.*

(2) *If an annihilator right ideal A contains a nonzero element of a member M of F, then $A \supset M$.*

(3) *Each member M of F contains (not necessarily different) elements a and b such that $0 \neq ab \in M$ and $Sa \neq 0$.*

Then S is a J_r-ring.

Proof. Let A be a left ideal of S, and suppose that $r(S,A) \neq 0$. Then $AM = 0$ for some $M \in F$ by (1). By (3) M contains a, b such that $0 \neq ab \in M$, $0 \neq Sa$. We shall show that $A \cap Sa = 0$. If not, $0 \neq ta \in A$ for some t. Since $Aa = 0$, $ta^2 = 0$, and hence $r(S, ta)$ contains $a \in M$. Thus, $taM = 0$ by (2), and so $tab = 0$. Now $r(S, l(S, ab))$ contains $ab \in M$. Hence $l(S, ab)M = 0$ by (2). Since $l(S, ab) \ni t$, we have $ta = 0$, a contradiction, completing the proof.

Remark. Let S be a ring with a nonzero idempotent in every nonzero annihilator right ideal. Then it is evident that S is a J_r-ring. The set of nonzero idempotents satisfies the conditions in Lemma 4.1.

LEMMA 4.2. *Let S be a semiprime ring, and let F be a family of nonzero right ideals satisfying the conditions* (1) *and* (2) *in Lemma* 4.1. *Then there exists a family G of left ideals which satisfies the following conditions:*

(1_l*) *Every nonzero left ideal contains a member of G.*

(2_l) *Let $0 \neq a$, $b \in N \in G$. Then $r(S,a) = r(S,b)$.*

Proof. Let G be the set of all those left ideals Sx that x is nonzero and belongs to a member of F.

Let A be a nonzero left ideal of S. Then $r(S, l(S,A))$ is nonzero, and contains a member B of F. Hence $l(S,A)B = 0$, and so $B \not\subset l(S,A)$ since $B^2 \neq 0$. Thus, $0 \neq BA \subset A \cap B$. Let $0 \neq y \in A \cap B$. Then $A \supset Sy \in G$.

Next, let $Sz \in G$, $0 \neq z \in C \in F$, and let uz be any nonzero element of Sz. Evidently $r(S, uz) \supset r(S, z)$. If $r(S, uz) \neq r(S, z)$, then $uzt = 0$ and $zt \neq 0$ for some t. Hence $0 \neq zt \in r(S, u) \cap C$, and so $r(S, u) \supset C$ by (2). Thus, $uz = 0$, a contradiction. Therefore $r(S, uz) = r(S, z)$. It follows from this that $r(S, a) = r(S, b)$ for any $0 \neq a$, $b \in Sz$.

THEOREM 4.3. *Let S be a semiprime ring. Suppose that there is given a family F of right ideals satisfying the conditions* (1) *and* (2) *in Lemma* 4.1. *Then S is a J-ring.*

Proof. Evidently F satisfies (3) too. Hence S is a J_l-ring by Lemma 4.1. By Lemma 4.2 there is a family G of left ideals satisfying (1_l*) and (2_l). (1_l*) implies the right-left symmetry of (1). It is easy to see that (2_l) is equivalent to the symmetry of (2). The symmetry of (3) is obviously fulfilled by G. Thus, by the right-left symmetry of Lemma 4.1 it follows that S is a J_r-ring, as desired.

COROLLARY 4.4. *Let S be a semiprime ring. Suppose that every nonzero annihilator right ideal contains a minimal annihilator right ideal. Then S is a J-ring.*

Proof. The set of minimal annihilator right ideals satisfies the conditions (1) and (2) in Lemma 4.1. Hence S is a J-ring by Theorem 4.3.

THEOREM 4.5. *Let S be a semiprime ring. Suppose that there is given a mapping v of the set L of left ideals of S into itself which satisfies the following conditions:*

(a) $v(0)=0$.

(b) $A \subset v(A)$ *and* $v(A)=v(v(A))$ *for any* $A \in L$.

(c) $v(A) \cap v(B)=v(A \cap B)$ *for any* $A, B \in L$.

(d) $v(Ax) \supset v(A)x$ *for any* $A \in L$ *and* $x \in S$.

We call a left ideal A v-closed if $A=v(A)$.

(e) *Every nonzero annihilator (two-sided) ideal contains a minimal nonzero v-closed left ideal.*

Then S is a J-ring.

Proof. Denote by F' the set of minimal nonzero v-closed left ideals of S.

(i) Every nonzero closed left ideal contains a member of F'. In fact, let A be a nonzero closed left ideal. Then $r(S, l(S, A))=l(S, l(S, A))(=B)$. By (e) B contains a member C of F'. If $CA=0$, then $C^2 \subset l(S, A) r(S, l(S, A))=0$, and so $C=0$, a contradiction. Hence $CA \neq 0$. Let $Cx \neq 0$ for $x \in A$. By (c) $v(Cx) \subset v(A)=A$. Suppose that $v(Cx) \supset D=v(D) \neq 0$. By (c) $v(Cx \cap D)=v(Cx) \cap D=D \neq 0$. Hence $Cx \cap D \neq 0$. Set $E=\{y \in C; yx \in D\}$. Then $E \neq 0$. By (d) $v(E)x \subset v(Ex) \subset v(D)=D$, and also $v(E) \subset v(C)=C$. Thus, $v(E)=E$. Since C is minimal, $C=E$, and hence $Cx \subset D$, and so $v(Cx) \subset v(D)=D$, which shows that $v(Cx) \in F'$, as desired.

(ii) Every annihilator left ideal is v-closed. In fact, any intersection

of v-closed left ideals is v-closed by (b) and (c). Hence it is enough to see that the left annihilator of an element x is v-closed. Set $A = l(S, x)$. By (d) $v(A)x \subset v(Ax) = v(0) = 0$, whence A is v-closed, as desired.

(iii) Combining (i) and (ii) we obtain that any nonzero annihilator left ideal contains a member of F'. If a nonzero annihilator left ideal A has a nonzero intersection with $B \in F'$, then A is v-closed by (ii), and so $A \cap B$ is also v-closed. Since B is minimal, $A \supset B$. Thus F' satisfies the right-left symmetry of the conditions (1) and (2) in Lemma 4.1, and therefore S is a J-ring by the symmetry of Theorem 4.3.

5. Let S be a J_l-ring. Then it is easily seen that the following three conditions for a left ideal A of S are equivalent:

(i) A is uniform.

(ii) For any $x \in A$ and any nonzero $y \in A$ there is a large left ideal B of S such that $Bx \subset Sy$.

(iii) For any $x \in A$ and any nonzero $y \in A$ there exists a left ideal B such that $Bx \subset Sy$ and $r(S, B) = 0$.

A ring S is called right faithful if $Sx = 0$, $x \in S$ imply $x = 0$. We say that a left ideal A of a right faithful ring S is strongly uniform if A satisfies the above condition (iii). A strongly uniform right ideal of a left faithful ring is defined symmetrically.

THEOREM 5.1. *Let S be a prime ring. Then the following conditions are equivalent:*

(i) *S is a J-ring with a uniform left ideal and a uniform right ideal.*

(ii) *S contains a strongly uniform left ideal and a uniform right ideal.*

(iii) *S has a strongly uniform left ideal and a strongly uniform right ideal.*

Proof. Every uniform left (right resp.) ideal of a J_l- (J_r resp.) ring is strongly uniform. Every strongly uniform left (right resp.) ideal of a right (left resp.) faithful ring is uniform. Hence the theorem directly follows from the following lemma which will be shown in the next section.

LEMMA 5.2. *If a prime ring S has a strongly uniform left ideal, then S is a J-ring.*

6. Let S be a right faithful ring. We denote by $N_l(S)$ or simply N_l

the set of left ideals A of S such that $r(S,A)=0$. Thus, $S\in N_l(S)$. We write $A\tilde{\rightarrow}B$ for left ideals A and B of S if for any element $x\in A$ we can find a left ideal $C\in N_l$ such that $Cx\subset B$.

LEMMA 6.1. *$A\tilde{\rightarrow}B\tilde{\rightarrow}C$ implies $A\tilde{\rightarrow}C$.*

Proof. Let $x\in A$. Then $Px\subset B$ for some $P\in N_l$. For any $p\in P$ we have a left ideal $Q_p\in N_l$ such that $Q_ppx\in C$. Set $R=\sum_{p\in P}Q_pp$. Then $Rx\subset C$. For any nonzero $y\in S$ there is $p\in P$ with $py\neq 0$, and hence $qpy\neq 0$ for some $q\in Q_p$. Since $qp\in R$, this shows that $R\in N_l$. Therefore $A\tilde{\rightarrow}C$, as desired.

We write $A\sim B$ if $A\tilde{\rightarrow}B$ and $B\tilde{\rightarrow}A$. Then, by Lemma 6.1 it is easily seen that the relation $\sim$ is an equivalence.

A left ideal A of S is called strongly large if $A\sim S$. (See [5; (1.6)], where the set of strongly large left ideals is denoted by S^{Δ}). For any left ideal A of S we denote by $d(A)$ $(=\Delta_S{}^S(A)$ in [5; p. 6]) the set of elements x such that $Bx\subset A$ for some strongly large left ideal B. Then $d(A)$ is a left ideal by [5; (3.1)].

LEMMA 6.2. *Every strongly large left ideal A of S belongs to $N_l(S)$.*

Proof. Let $0\neq x\in S$. Since S is right faithful, $yx\neq 0$ for some $y\in S$. There exists a left ideal $B\in N_l$ with $By\subset A$. Now we can find an element $b\in B$ such that $byx\neq 0$. Hence Ax contains $byx\neq 0$, as desired.

LEMMA 6.3. *d satisfies the conditions* (a), (b), (c) *and* (d) *in Theorem* 4.5.

Proof. If $d(0)\ni x$, then $Ax=0$ for some strongly large left ideal A. Hence $x=0$ by Lemma 6.2. This shows that d fulfilled (a). (b) follows from [5; (3.1) and (3.6)]. (c) and (d) coincide with [5; (3.2) and (3.4)].

LEMMA 6.4. *Let A and B be left ideals of S. $A\sim B$ if and only if $d(A)=d(B)$.*

Proof. Suppose that $A\sim B$. Set $C_x=\{y\in S: yx\in B\}$ for any $x\in A$. Let $z\in S$. Then $zx\in A$, $C_{zx}\in N_l$, and $C_{zx}z\subset C_x$. Hence C_x is strongly large. This shows that $A\subset d(B)$, and so $d(A)\subset d(B)$. Similarly $d(A)\supset d(B)$. Therefore $d(A)=d(B)$, as desired.

LEMMA 6.5. *Let A be a left ideal of S. Then $d(A)$ is a minimal nonzero d-closed left ideal if and only if A is strongly uniform.*

Proof. Suppose that $d(A)$ is a minimal nonzero d-closed left ideal, and let $0 \neq x, y \in A$. Then $d(A) = d(Sy)$ since $A \supset Sy \neq 0$. Denote by B the left ideal generated by x. We have $d(A) = d(B)$, too. Hence $d(B) = d(Sy)$, and so $B \sim Sy$ by Lemma 6.4. Thus, $Cx \subset Sy$ for some $C \in N_l$, which shows that A is strongly uniform.

Conversely, let A be a strongly uniform left ideal. Let $0 \neq a \in A$. Then $A \overset{\sim}{\rightarrow} Sa$, and hence $A \sim Sa$. Thus $d(A) = d(Sa)$ by Lemma 6.4. This implies that A is a minimal nonzero d-closed left ideal, as desired.

Proof of Lemma 5.2. We have seen in Lemma 6.3 that d satisfies (a), (b), (c) and (d) in Theorem 4.5. Let A be a nonzero annihilator ideal of S. Then, since S is prime, $A = S$. Hence A contains a strongly uniform left ideal B by assumption. $d(B)$ is a minimal nonzero d-closed left ideal by Lemma 6.5. Therefore d satisfies the condition (e) too. By Theorem 4.5 S is a J-ring, completing the proof.

7. Finally we note the following

THEOREM 7.1. *Let S be a prime ring with a uniform left ideal and a uniform right ideal. Then the following conditions are equivalent:*

(i) *S is a J-ring.*

(ii) *Every nonzero annihilator left ideal contains a minimal nonzero annihilator left ideal.*

Proof. By the right-left symmetry of Corollary 4.4 (ii) implies (i).

Let S be a J-ring with a uniform left ideal and a uniform right ideal, and let A be an annihilator left ideal $\neq 0$. As is known any annihilator left ideal of a J_r-ring is closed (in the sense of Section 3). Now the lattice $L(\bar{S}_t)$ of closed left ideals of $\bar{S}_t$ is isomorphic to the lattice $L(S)$ of closed left ideals of S by the correspondence $W(\in L(\bar{S}_t)) \rightarrow W \cap S$. Hence there is $B \in L(\bar{S}_t)$ such that $B \cap S = A$. Since $\bar{S}_t$ is a primitive ring with non-zero socle, B contains $\bar{S}_t e$ with a primitive idempotent e of $\bar{S}_t$. $\bar{S}_t e$ is a minimal nonzero annihilator left ideal of $\bar{S}_t$, and hence $\bar{S}_t e \cap S$ $(= C)$ is also a minimal nonzero annihilator left ideal of S in view of the next lemma. Since $C \subset A$, this proves the theorem.

LEMMA 7.2. *Let R be a right quotient ring of a J_r-ring S, and D an annihilator left ideal of R. Then $D \cap S$ is an annihilator left ideal of S. In fact, $D \cap S = l(S, r(S, D))$.*

Proof. Set $E = r(R, D)$. Evidently

$$D \cap S \subset l(S, E \cap S).$$

Let $x \in l(S, E \cap S)$. Then $E \cap S \subset r(S, x) = r(R, x) \cap S$, where both E and $r(E, x)$ are closed right ideals of R. By virtue of the isomorphism between the lattices of closed right ideals of R and S, it follows that $E \subset r(R, x)$. Hence $x \in l(R, E) \cap S = D \cap S$. This shows that $D \cap S = l(S, E \cap S)$, as desired.

As a consequence of this lemma we obtain the following, which may be interesting in connection with [6; (K'_1) in Theorem 2.2].

PROPOSITION 7.3. *Let S be a prime J-ring with a uniform left ideal and a uniform right ideal, and let $L(S)$ be the lattice of closed left ideals of S. Then every finite dimensional element of $L(S)$ is an annihilator left ideal.*

Proof. If A is finite dimensional in $L(S)$, then $A = B \cap S$ for some finite dimensional B in $L(\bar{S}_t)$. B is generated by an idempotent in $\bar{S}_t$ by Proposition 3.3, and hence it is an annihilator left ideal. Therefore A is also an annihilator left ideal of S, as desired.

Another consequence of Lemma 7.2 is the following

PROPOSITION 7.4. *Let R be a two-sided quotient ring of a J-ring S, and denote the lattices of annihilator left ideals of R and S by $L^*(R)$ and $L^*(S)$ respectively. Then the correspondence $W \ (\in L^*(R)) \to W \cap S$ gives an isomorphism of $L^*(R)$ and $L^*(S)$.*

Proof. Denote the correspondence by v. v is a mapping of $L^*(R)$ into $L^*(S)$ by Lemma 7.2. Let $A \in L^*(S)$, and let $A = l(S, B)$. Then $A = l(R, B) \cap S$, whence v is onto. Moreover v is one-to-one since it is a restriction of the isomorphism between the lattices of closed left ideals of R and S, as desired.

8. It is easy to find a J-ring with a uniform left ideal and with no uniform right ideal. In fact, let S be a left Ore domain, and suppose that it is not a right Ore domain. Then $\bar{S}_l$ is a division ring, the (classical) left quotient division ring of S. S itself is uniform as a left ideal. If S contains a uniform right ideal, then $\bar{S}_t$ is a primitive ring with nonzero socle by Theorem 3.1. Since $\bar{S}_l \supset \bar{S}_t \supset S$, $\bar{S}_t$ is also a division ring, therefore $\bar{S}_l = \bar{S}_t = \bar{S}_r$. This implies that S is a uniform right ideal of S; in other words, S is a right Ore domain, a contradiction.

THE UNIVERSITY OF ROCHESTER.

REFERENCES.

[1] A. W. Goldie, "The structure of prime rings under ascending chain conditions," *Proceedings of the London Mathematical Society*, vol. 8 (1958), pp. 589-608.

[2] N. Jacobson, *Structure of rings*, American Mathematical Society Colloquium Publications, vol. 37 (1956).

[3] R. E. Johnson, "The extended centralizer of a ring over a module," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 891-895.

[4] ———, "Representations of prime rings," *Transactions of the American Mathematical Society*, vol. 74 (1953), pp. 351-357.

[5] Yuzo Utomi, "On quotient rings," *Osaka Mathematical Journal*, vol. 8 (1956), pp. 1-18.

[6] ———, "On rings of which any one-sided quotient rings are two-sided," *Proceedings of the American Mathematical Society*, vol. 14 (1963), pp. 141-147.

ON THE INDEX OF ELLIPTIC OPERATORS ON CLOSED SURFACES.*[1]

By WALTER KOPPELMAN.

The index of a linear mapping is the difference between the dimension of its null space and the codimension of its range. Many results concerning the index of linear elliptic operators on two-dimensional domains have been known for some time. In a recent article by I. M. Gelfand [12], a systematic study of the index of general elliptic problems as a homotopy invariant was suggested. Subsequently articles by Volpert, Dynin, and Agranovič [1, 2, 8, 9, 23, 24] have contributed substantially to the understanding of the nature of this problem. The author has recently learned that M. F. Atiyah and I. M. Singer have obtained a formula for the index of very general elliptic problems on compact manifolds of any dimension. Unfortunately, these results are unavailable to him at this time.[2]

The present paper was inspired by two announcements of Volpert [23, 24], concerning elliptic problems on the sphere, S^2. The second of these contains an especially simple formula for the index in terms of the degree of a related mapping. In this work, we show that this formula remains valid for any closed oriented surface. Furthermore, we make a detailed study of the structure of elliptic operators of first order on such a surface.

For the convenience of the reader, we have included an exposition of known results in §§ 1-3. With slight modifications, these remain valid for compact manifolds of arbitrary finite dimension. In § 4, we define a homomorphism κ, which sends the group of continuous mappings of the cotangent sphere bundle $\mathfrak{B}$ of the surface into the general linear group $GL(n, C)$, into the additive group of integers. The value of κ on a mapping $\mathfrak{A}: \mathfrak{B} \to GL(n, C)$ of class C^∞ is just the index of a system of singular integral operators, whose symbol is the function $\mathfrak{A}$. § 5 is devoted to a study of the general structure of elliptic operators of first order on closed surfaces. In § 6, we define the degree $l(\mathfrak{A})$ of a mapping $\mathfrak{A}: \mathfrak{B} \to GL(n, C)$. This is just the degree of the

* Received April 1, 1963.

[1] This research was supported by the Air Force Office of Scientific Research.

[2] *Added in proof.* The results of Atiyah and Singer have been announced in the *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 422-433.

mapping $\mathcal{B} \to U(n) \to U(n)/U(1)$ defined by the unitary part U of the polar decomposition $({}^t\bar{\mathfrak{A}}\mathfrak{A})^{\frac{1}{2}}U$ of $\mathfrak{A}$. Finally, in § 7, we show that, after a proper selection of the generator of the third homology group, $H_3(U(n)/U(1); Z)$, the equality $\kappa(\mathfrak{A}) = l(\mathfrak{A})$ is always valid. In particular, one obtains the result that the index of an $n \times n$ elliptic system of order m (see § 1) is equal to the degree $l(\mathfrak{A})$ of the mapping $\mathfrak{A}: \mathcal{B} \to GL(n, C)$ furnished by its characteristic matrix function $\mathfrak{A}$.

The author is indebted to Professor C. T. Yang for several helpful discussions about some of the topological aspects of this paper.

1. General remarks concerning linear elliptic differential operators. We shall be concerned with an $n \times n$ system of linear partial differential operators on a closed orientable surface $\mathcal{S}$ of class C^∞. Such a system has the form

$$Lu \equiv \sum A^{i_1 i_2 \cdots i_m} \frac{\partial^m u}{\partial x^{i_1} \partial x^{i_2} \cdots \partial x^{i_m}} + \text{lower order terms}, \tag{1.1}$$

relative to a local coördinate system (x^1, x^2) on $\mathcal{S}$. Here $A^{i_1 i_2 \cdots i_m}$ is an $n \times n$ matrix, whose entries are the $(i_1 i_2 \cdots i_m)$-coefficients of given contravariant C^∞ tensor fields of rank m, in terms of the natural coördinate base. The operator acts on $n \times 1$ vector functions u, and its range also consists of such functions.

For a covariant vector with coefficients (ξ_1, ξ_2) relative to the natural coördinate dual base, we consider the characteristic matrix function

$$\mathfrak{A} \equiv \sum A^{i_1 i_2 \cdots i_m} \xi_{i_1} \xi_{i_2} \cdots \xi_{i_m}. \tag{1.2}$$

Clearly, this defines a function on the cotangent bundle of the surface $\mathcal{S}$. The operator (1.1) is said to be *elliptic* on $\mathcal{S}$, if, for every point of the cotangent bundle which does not belong to the zero cross section,

$$\det \mathfrak{A} \neq 0. \tag{1.3}$$

Consequently, an elliptic differential operator provides a mapping of the nonzero part of the cotangent bundle into the general linear group, $GL(n, C)$.

2. The theory of singular integral operators. We shall also consider $n \times n$ systems of singular integral operators on $\mathcal{S}$. Let us first discuss a single singular integral operator on R^2. We consider a function $h(x, z)$ on $R^2 \times (R^2 - \{0\})$ of class C^∞, with

$$h(x, \lambda z) = \lambda^{-2} h(x, z), \text{ for all } \lambda > 0, \tag{2.1}$$

and

$$\int_{|z|=1} h(x,z)\,d\sigma = 0. \tag{2.2}$$

A singular integral operator on $L^2(R^2)$ then takes the form

$$(Hf)(x) \equiv a(x)f(x) + \lim_{\epsilon\to 0}\int_{|x-y|>\epsilon} h(x, x-y)f(y)\,dy, \tag{2.3}$$

where $a(x)$ is a bounded C^∞ function on R^2, and the limit is taken in terms of the norm of $L^2(R^2)$. According to a theorem of Calderon and Zygmund [5, p. 909],

$$\| Hf \| \leqq \| f \| \cdot C \sup_{x,|z|=1} (|a(x)| + |h(x,z)|), \tag{2.4}$$

where C is a constant which is independent of f.

We now pass to the surface $\mathcal{S}$. On $\mathcal{S}$ we consider a partition of unity $\{\phi_\alpha\}$ of class C^∞, where the support of ϕ_α lies in a single coördinate neighborhood with local coördinates $(x^1{}_\alpha, x^2{}_\alpha)$. We now consider for each $f \in C^\infty(\mathcal{S})$, the norms

$$\| f \|_r = \Big(\sum_{|p|\leqq r}\sum_\alpha \| \phi_\alpha^{\frac{1}{2}}(\partial/\partial x_\alpha)^p f \|^2\Big)^{\frac{1}{2}}, \tag{2.5}$$

where $\| \;\|$ denotes the norm of $L^2(R^2)$. The completion of $C^\infty(\mathcal{S})$ with respect to the norm $\| \;\|_r$, will be denoted by $L^2{}_r(\mathcal{S})$. It is known that the natural map, $L^2{}_r(\mathcal{S}) \to L^2{}_{r-1}(\mathcal{S})$ is a compact linear imbedding of $L^2{}_r(\mathcal{S})$ in $L^2{}_{r-1}(\mathcal{S})$ [19].

Definition 2.1. An operator R is called a *smoothing operator* if both R and R^* are bounded linear transformations of $L^2{}_r(\mathcal{S})$ into $L^2{}_{r+1}(\mathcal{S})$ for any integer r, $r \geqq 0$. (R^* is defined relative to $L^2{}_0(\mathcal{S})$.)

Definition 2.2. An operator S on $L^2{}_0(\mathcal{S})$ will be called a *singular integral operator* on $\mathcal{S}$, if

(i) for $\phi, \psi \in C^\infty(\mathcal{S})$ with disjoint supports, $\phi S \psi$ is a smoothing operator, and

(ii) for each pair $\phi, \psi \in C^\infty(\mathcal{S})$ with support in a common coördinate system with coördinates $x = (x^1, x^2)$, $\phi S\psi = \phi H\psi + R$, where H is a Euclidean singular integral operator of the form (2.3), while R is a smoothing operator.

Definition 2.3. The *symbol* of a singular integral operator S on $\mathcal{S}$ is the function on the contangent bundle of $\mathcal{S}$ defined by

$$\sigma(S)(p, \sum \xi_j dx^j) = a(x(p)) + \lim_{\epsilon\to 0}\int_{\epsilon^{-1}>|\eta|>\epsilon} \exp(i \sum_j \xi_j \eta_j) h(x(p), \eta)\,d\eta,$$

where $a(x), h(x,\eta)$ are the functions which are associated with the operator H of part (ii) of the preceding definition, and p lies in the intersection of the supports of ϕ and ψ.

It can be shown that $\sigma(S)$ is well-defined [20]. Observe that the symbol is positive homogeneous of degree zero in the ξ-variables.

Let $\mathcal{B}$ denote the cotangent sphere bundle of $\mathcal{S}$. Then each function of the space $C^\infty(\mathcal{B})$ is the symbol of a singular integral operator S on $\mathcal{S}$ and conversely, for each singular integral operator S on $\mathcal{S}$, the restriction of the symbol $\sigma(S)$ to $\mathcal{B}$ belongs to $C^\infty(\mathcal{B})$. If S_1, S_2 are singular integral operators on $\mathcal{S}$ and c_1, c_2 are constants, then $\sigma(c_1S_1+c_2S_2)=c_1\sigma(S_1)+c_2\sigma(S_2)$. If S is a singular integral operator on $\mathcal{S}$, such that $\sigma(S_1)\sigma(S_2)=\sigma(S)$, then $S-S_1S_2$ is a smoothing operator. Relative to a fixed Riemannian metric on $\mathcal{S}$ of class C^∞, we form the Hilbert space $L^2(\mathcal{S})$ defined by the volume element of the metric. If S is a singular integral operator on $\mathcal{S}$, then S^*, the adjoint of S relative to $L^2(\mathcal{S})$, is also a singular integral operator on $\mathcal{S}$, and $\sigma(S^*)=\overline{\sigma(S)}$. Finally, we note that for every function $F\in C^\infty(\mathcal{B})$, it is possible to construct a singular integral operator S_F, whose norm satisfies

$$\|S_F\|\leqq C\Big(\sum_{|p|\leqq 2\cdot\dim\mathcal{S}}\sup|D^pF|\Big), \tag{2.6}$$

where C is independent of F. (The reader should have no difficulty in giving a proper interpretation to the right side of (2.6).) [20].

In terms of the Riemannian metric on $\mathcal{S}$, we consider the harmonic operator Δ and its extension to the space $L^2{}_2(\mathcal{S})$ in the usual manner. This extension, which we still denote by Δ, is self-adjoint relative to the Hilbert space $L^2(\mathcal{S})$. For a suitable positive constant c, it is possible to show [20] that

$$(c\cdot 1-\Delta)^{\frac{1}{2}}=(c\cdot 1-\Delta)J \tag{2.7}$$

where J is a smoothing operator which commutes with $(c\cdot 1-\Delta)$ and for which the identity

$$J^2=(c\cdot 1-\Delta)^{-1} \tag{2.8}$$

holds. From these formulae, we see that $(c\cdot 1-\Delta)^{\frac{1}{2}}$ is a bounded mapping of $L^2{}_r(\mathcal{S})$ into $L^2{}_{r-1}(\mathcal{S})$. The mapping J is a two-sided inverse. Hence $(c\cdot 1-\Delta)^{\frac{1}{2}}$ is a topological isomorphism of $L^2{}_r(\mathcal{S})$ onto $L^2{}_{r-1}(\mathcal{S})$.

If $V_1,\cdots,V_m$ are given C^∞ vector fields on $\mathcal{S}$, then

$$V_1V_2\cdots V_m=S(c\cdot 1-\Delta)^{m/2}, \tag{2.9}$$

where S is a singular integral operator on $\mathcal{S}$, with the symbol

$$(2.10) \qquad (-i)^m \prod_{k=1}^{m} \langle V_k, \xi / |\xi| \rangle.$$

Here $\langle\ \rangle$ gives the pairing of the tangent and cotangent vectors of $\mathcal{S}$ into the field of real numbers [20].

If we consider a single differential operator

$$(2.11) \qquad Lu \equiv \sum a^{i_1 i_2 \cdots i_m} \frac{\partial^m u}{\partial x^{i_1} \partial x^{i_2} \cdots \partial x^{i_m}} + \text{terms of lower order},$$

on $\mathcal{S}$, then we can write $L = \sum_\alpha \phi_\alpha L$ by means of our partition of unity. Each operator $\phi_\alpha a^{i_1 \cdots i_m} \dfrac{\partial^m}{\partial x^{i_1} \cdots \partial x^{i_m}}$ is then a product of m vector fields. In fact, consider a function $\psi_\alpha \in C^\infty(\mathcal{S})$, whose values are equal to 1 in a neighborhood of the support of ϕ_α, but whose support lies in the coördinate neighborhood associated with ϕ_α. Then we can take

$$V_1 = \phi_\alpha a^{i_1 \cdots i_m} \frac{\partial}{\partial x^{i_1}}, V_2 = \psi_\alpha \frac{\partial}{\partial x^{i_2}}, \cdots, V_m = \psi_\alpha \frac{\partial}{\partial x^{i_m}}.$$

Of course,

$$(2.12) \quad \langle V_1, \xi/|\xi| \rangle = \phi_\alpha a^{i_1 \cdots i_m} \xi_{i_1}/|\xi|, \cdots, \langle V_m, \xi/|\xi| \rangle = \psi_\alpha \xi_{i_m}/|\xi|,$$

and

$$(2.13) \qquad V_1 V_2 \cdots V_m u = \phi_\alpha a^{i_1 \cdots i_m} \frac{\partial^m u}{\partial x^{i_1} \cdots \partial x^{i_m}}.$$

The corresponding singular integral operator $S_{\alpha, i_1, \cdots, i_m}$ has the symbol

$$(2.14) \qquad (-i)^m \phi_\alpha a^{i_1 \cdots i_m} \xi_{i_1} \cdots \xi_{i_m} / |\xi|^m.$$

We now set $S_\alpha = \sum_{i_1, \cdots, i_m} S_{\alpha, i_1, \cdots, i_m}$ and $S = \sum_\alpha S_\alpha$. Then

$$\phi_\alpha L - S_\alpha (c \cdot 1 - \Delta)^{m/2}$$

involves only terms of order less than m. These clearly give a compact mapping of $L^2{}_m(\mathcal{S})$ into $L^2{}_0(\mathcal{S})$. Consequently

$$(2.15) \qquad L - S(c \cdot 1 - \Delta)^{m/2}$$

is a compact mapping of $L^2{}_m(\mathcal{S})$ into $L^2{}_0(\mathcal{S})$.

Let us now turn to the discussion of systems. Suppose that L is the linear operator of the expression (1.1). The operator L may then be written as a matrix of differential operators

$$(2.16) \qquad L = (L_{ij}), i, j = 1, \cdots, n,$$

so that $Lu = (\sum_{j=1}^{n} L_{ij}u_j)$, $i = 1, \cdots, n$. We assume that the order of L is m. This means that the order of each L_{ij} is less than or equal to m. By applying the preceding discussion to the m-th order terms of the operators L_{ij}, we see that

$$L_{ij} = S_{ij}(c\cdot 1 - \Delta)^{m/2} + K_{ij}, \tag{2.17}$$

where K_{ij} is a compact linear mapping of $L^2{}_m(\mathcal{S})$ into $L^2{}_0(\mathcal{S})$. We set

$$S = (S_{ij}), K = (K_{ij}), i, j = 1, \cdots, n, \tag{2.18}$$

$$\Lambda = (\delta_{ij}(c\cdot 1 - \Delta)^{\frac{1}{2}}), i, j = 1, \cdots, n, \tag{2.19}$$

so that

$$L = S\Lambda^m + K. \tag{2.20}$$

For the vector function $u = (u_i)$, $i = 1, \cdots, n$, we introduce the norms

$$(\| u \|_{r,n})^2 = \sum_{i=1}^{n} (\| u_i \|_r)^2, \tag{2.21}$$

and we speak of the corresponding spaces $L^2{}_{r,n}(\mathcal{S})$. Thus K is a compact linear mapping of $L^2{}_{m,n}(\mathcal{S})$ into $L^2{}_{0,n}(\mathcal{S})$, while the diagonal operator Λ^m is a topological isomorphism of $L^2{}_{m,n}(\mathcal{S})$ onto $L^2{}_{0,n}(\mathcal{S})$.

A system $S = (S_{ij})$, $i, j = 1, \cdots, n$ of singular integral operators is said to be *elliptic* on $\mathcal{S}$, if the $n \times n$ matrix of symbols, $\sigma(S) = (\sigma(S_{ij}))$, which we shall call the symbol of S, is non-singular at every point of the cotangent sphere bundle $\mathfrak{B}$ of $\mathcal{S}$. It is quite clear that, if L is an elliptic system of differential operators, then the corresponding system S of singular integral operators in the construction above, will also be elliptic.

The Hilbert space $L^2(\mathcal{S})$, which is formed through the volume element of the Riemannian metric, is topologically equivalent to the space $L^2{}_0(\mathcal{S})$. For $n \times 1$ vector functions $u = (u_i), v = (v_i)$, $i = 1, \cdots, n$, we consider the scalar product

$$(u, v)_n = \sum_{i=1}^{n} (u_i, v_i), \tag{2.22}$$

where (u_i, v_i) is the scalar product in $L^2(\mathcal{S})$. We thus obtain a Hilbert space $L^2{}_{,n}(\mathcal{S})$ which is topologically equivalent to $L^2{}_{0,n}(\mathcal{S})$. If S is an $n \times n$ system of singular integral operators, then we have

$$(Su, v)_n = \sum_{i=1}^{n} \sum_{j=1}^{n} (S_{ij}u_j, v_i) = \sum_{j=1}^{n} \sum_{i=1}^{n} (u_j, S_{ij}{}^*v_i) = (u, S^*v)_n.$$

We have previously noted that S_{ij}^* is a singular integral operator with $\sigma(S_{ij}^*) = \overline{\sigma(S_{ij})}$. Consequently, if $\sigma(S)$ is the symbol of the system S, then $\sigma(S^*) = {}^t\overline{\sigma(S)}$. Thus, if S is elliptic, then S^* is also elliptic.

Calderon and Zygmund [4] have shown that if H is a Euclidean singular integral operator with constant coefficients, i. e., the functions $a(x), h(x,z)$ in (2.1)-(2.3) are independent of x, then the representation of the operator H after a Fourier transformation of $L^2(R^2)$ is given by

$$(2.23) \qquad (\widehat{Hf})(\xi) = \sigma(H)(\xi)\cdot \hat{f}(\xi),$$

where $\sigma(H)$ is the symbol of H. This functional calculus together with the estimate (2.4) enables one to show that if S is an elliptic system of singular integral operators on $\mathcal{S}$, then an estimate of the form

$$(2.24) \qquad \| u \|_{2,n} \leqq C(\| S\Lambda^2 u \|_{0,n} + \| u \|_{0,n})$$

is valid for all $u \in L^2{}_{2,n}(\mathcal{S})$ [2, 8]. The constant C does not depend on u. Hence, on the null space of $S\Lambda^2$, which is a closed linear subspace of $L^2{}_{2,n}(\mathcal{S})$, we have the estimate

$$(2.25) \qquad \| u \|_{2,n} \leqq C \| u \|_{0,n} \leqq C' \| u \|_{2,n}.$$

It follows that the unit sphere of the null space of $S\Lambda^2$ is compact, so that the space is of finite dimension [7, p. 245]. Consequently we also see that

$$(2.26) \qquad \dim \ker S < \infty.$$

Since S^* is also elliptic in the present case, we have $\dim \ker S^* < \infty$. Next, for those $u \in L^2{}_{2,n}(\mathcal{S})$ which satisfy $(u, \ker S\Lambda^2)_n = 0$, one can obtain the estimate [cf. 17]

$$(2.27) \qquad \| u \|_{2,n} \leqq C'' \| S\Lambda^2 u \|_{0,n}.$$

From (2.27) it follows that the range of $S\Lambda^2$ is closed. Consequently the range of S is closed. The orthogonal complement of the range of S is just the kernel of S^*. This orthogonal complement is then of finite dimension. Since the range of S is closed while both the kernel and the cokernel of S are of finite dimension, we see that an elliptic system of singular integral operators furnishes a *Φ-operator* on the space $L^2{}_{,n}(\mathcal{S})$. [13]

3. **Remarks concerning the index.** Let B_1, B_2 be two Banach spaces and let $L\colon B_1 \to B_2$ be a bounded linear operator which is also a Φ-operator, i. e., L has a closed range and the spaces $\ker L$, B_2/LB_1, are of finite dimension. The index of the operator L is the integer

$$\operatorname{ind} L = \dim \ker L - \dim B_2/LB_1. \tag{3.1}$$

The following results are known [13].

THEOREM 3.1. *If K is a compact linear operator from B_1 into B_2, while L is a Φ-operator, then $L+K$ is a Φ-operator, and*

$$\operatorname{ind}(L+K) = \operatorname{ind} L.$$

THEOREM 3.2. *For a given Φ-operator L, there exists a positive number ρ_L, such that for every bounded linear operator $M: B_1 \to B_2$ of norm $\| M \| < \rho_L$, the operator $L+M$ is a Φ-operator, and*

$$\operatorname{ind}(L+M) = \operatorname{ind} L.$$

THEOREM 3.3. *If L_1, L_2 are bounded Φ-operators on a Banach space B, then L_1L_2 is a bounded Φ-operator on B, and*

$$\operatorname{ind}(L_1L_2) = \operatorname{ind} L_1 + \operatorname{ind} L_2.$$

From these theorems, the decomposition (2.20) and the remarks after (2.21), it follows that if L is the operator corresponding to an $n \times n$ elliptic system, while S is the elliptic system of singular integral operators as in (2.20), then

$$\operatorname{ind} L = \operatorname{ind} S. \tag{3.2}$$

If S_1, S_2 are elliptic singular systems with $\sigma(S_1) = \sigma(S_2)$, then the discussion following Definition 2.3 shows that $S_1 - S_2$ is a smoothing operator (hence compact), so that

$$\operatorname{ind} S_1 = \operatorname{ind} S_2 \tag{3.3}$$

By combining the preceding result with (2.6) and Theorem 3.2, we see that if S is an elliptic singular system, then there exists a positive number ρ_S, such that for any singular system M with

$$\sum_{|p| \leqq 4} \sup | D^p \sigma(M) | < \rho_S,$$

$S+M$ is elliptic, and

$$\operatorname{ind}(S+M) = \operatorname{ind} S. \tag{3.4}$$

R. T. Seeley [21] has shown that if S is a single elliptic singular operator (with $n=1$), then

$$\operatorname{ind} S = 0. \tag{3.5}$$

4. The index as a homotopy invariant. Corresponding to each continuous mapping

$$(4.1) \qquad \mathcal{A}\colon \mathcal{B} \to GL(n, C)$$

of the cotangent sphere bundle $\mathcal{B}$ into the general linear group $GL(n, C)$, we shall consider an integer $\kappa(\mathcal{A})$ defined as follows. Suppose that the mapping $\mathcal{A}$ is homotopic to a mapping

$$F\colon \mathcal{B} \to GL(n, C)$$

of class C^∞. Then, according to the discussion after Definition 2.3, $F = \sigma(S)$, for an elliptic singular system S on $\mathcal{S}$. We set

$$(4.2) \qquad \kappa(\mathcal{A}) = \operatorname{ind} S.$$

In order to show that κ is well-defined, we must show that if F' is another C^∞ mapping, with $\mathcal{A} \sim F' = \sigma(S')$, then $\operatorname{ind} S = \operatorname{ind} S'$. Certainly we have $F \sim F'$. This means that there is a mapping

$$(4.3) \qquad F_t\colon \mathcal{B} \times I \to GL(n, C), \qquad (I = [0, 1])$$

such that $F_0 = F, F_1 = F'$. Let us select a sequence of points

$$0 = t_0 < t_1 < \cdots < t_\nu = 1,$$

such that $F_{t_{k+1}} = F_{t_k} e^{Q_k}$, where Q_k is a continuous matrix-valued function on $\mathcal{B}$. Next, for each $k \neq 0, \nu$, we select C^∞ mappings

$$G_k\colon \mathcal{B} \to GL(n, C),$$

such that $G_{k+1} = G_k e^{R_k}$, $G_1 = F e^{R_0}$, $F' = G_{\nu-1} e^{R_{\nu-1}}$. All R_k will be of class C^∞. Then,

$$(4.4) \qquad F' = F e^{R_0} e^{R_1} \cdots e^{R_{\nu-1}}.$$

Observe that for each k, e^{tR_k}, $0 \leqq t \leqq 1$, provides a smooth homotopy between e^{R_k} and the constant mapping 1. Hence, according to the remarks leading to (3.4), $\kappa(e^{R_k}) = 0$. By Theorem 3.3,

$$(4.5) \qquad \kappa(F') = \kappa(F e^{R_0} e^{R_1} \cdots e^{R_{\nu-1}}) = \kappa(F) + \sum_{k=0}^{\nu-1} \kappa(e^{R_k}) = \kappa(F).$$

Consequently $\kappa(\mathcal{A})$ is well-defined and depends only on the homotopy class of the mapping $\mathcal{A}$. Further, if $\mathcal{A}_1 \sim F$, $\mathcal{A}_2 \sim G$, where F, G are C^∞ mappings, and if

$$(4.6) \qquad \mathcal{A}_1 \mathcal{A}_2\colon \mathcal{B} \to GL(n, C)$$

is the mapping defined by matrix multiplication, then we clearly have $\mathcal{A}_1\mathcal{A}_2 \sim FG$, and

$$(4.7) \qquad \kappa(\mathcal{A}_1\mathcal{A}_2) = \kappa(FG) = \kappa(F) + \kappa(G) = \kappa(\mathcal{A}_1) + \kappa(\mathcal{A}_2).$$

5. The structure of first order elliptic operators on closed surfaces.

A. *The existence of operators.*

The local form of a first order operator is, of course,

$$(5.1) \qquad A^1 \frac{\partial u}{\partial x^1} + A^2 \frac{\partial u}{\partial x^2} + \text{lower order terms.}$$

The ellipticity condition states that

$$(5.2) \qquad \det(A^1\xi_1 + A^2\xi_2) \neq 0 \text{ for } (\xi_1)^2 + (\xi_2)^2 \neq 0.$$

In particular, $\det A^j \neq 0$, $j = 1, 2$. We shall be especially interested in the roots λ of the equation

$$(5.3) \qquad \det(A^1 - \lambda A^2) = 0.$$

Under the hypothesis of ellipticity, these roots can never be real. Further, because $\det A^2 \neq 0$, (5.3) is equivalent to

$$(5.4) \qquad \det(A - \lambda \cdot 1) = 0,$$

where $A = (A^2)^{-1}A^1$.

On our closed oriented surface $\mathcal{S}$, we now select an arbitrary conformal structure which is compatible with the C^∞ structure on $\mathcal{S}$ and its orientation [3, 6]. This conformal structure will remain fixed during the rest of the discussion. We now consider a small neighborhood D on $\mathcal{S}$, which is conformally equivalent to a disc. On D, we introduce a local uniformizer $|z| < 1$. We denote by D_ρ the region corresponding to $|z| < \rho$, where $0 < \rho < 1$. On $D - D_\rho$, we introduce the local uniformizer $\zeta = 1/z$. Then $d\zeta/dz = -z^{-2}$, and

$$(5.5) \qquad \tfrac{1}{2\pi}[\arg \frac{d\zeta}{dz}]_{\dot{D}_\rho} = -2,$$

where the boundary $\dot{D}_\rho$ of D_ρ is assumed to have positive orientation relative to D_ρ. We suppose that we have a first order elliptic operator defined on $\mathcal{S} - D_\rho$. With $\zeta = \xi^1 + i\xi^2$, $z = x^1 + ix^2$, and

$$(5.6) \qquad \hat{A} = (-A \frac{\partial x^1}{\partial \xi^2} + 1 \cdot \frac{\partial x^1}{\partial \xi^1})^{-1}(A \frac{\partial x^1}{\partial \xi^1} + 1 \cdot \frac{\partial x^1}{\partial \xi^2}),$$

(this is just the transformation law for $A = (A^2)^{-1}A^1$, when the Cauchy-Riemann equations are used), we now prove

LEMMA 5.1. $[\arg\det \hat{A}]_{\dot{D}_\rho} = 0$.

Proof. If $\hat{\lambda}_1, \cdots, \hat{\lambda}_n$ denote the eigenvalues of $\hat{A}$, then

$$\det \hat{A} = \prod_{j=1}^{n} \hat{\lambda}_j.$$

The eigenvalues are continuous on $\dot{D}_\rho$, and never real. Hence $[\arg \hat{\lambda}_j]_{\dot{D}_\rho} = 0$, and

$$[\arg\det \hat{A}]_{\dot{D}_\rho} = \sum_{j=1}^{n} [\arg \hat{\lambda}_j]_{\dot{D}_\rho} = 0.$$

LEMMA 5.2. *Let the matrix A possess P eigenvalues with positive imaginary part, N eigenvalues with negative imaginary part. Then*

$$\tfrac{1}{2\pi}[\arg\det(A\frac{\partial x^1}{\partial \xi^1} + 1\cdot\frac{\partial x^1}{\partial \xi^2})]_{\dot{D}_\rho} = 2(P-N).$$

Proof. Let $\lambda_1, \cdots, \lambda_n$ denote the eigenvalues of A. Then

$$\det(A\frac{\partial x^1}{\partial \xi^1} + 1\cdot\frac{\partial x^1}{\partial \xi^2}) = \prod_{j=1}^{n} (\lambda_j \frac{\partial x^1}{\partial \xi^1} + \frac{\partial x^1}{\partial \xi^2}).$$

The linear homotopies

$$(1-t)(\lambda_j\frac{\partial x^1}{\partial \xi^1} + \frac{\partial x^1}{\partial \xi^2}) + t(\operatorname{sgn}(\operatorname{Im}\lambda_j)\cdot i\frac{\partial x^1}{\partial \xi^1} + \frac{\partial x^1}{\partial \xi^2}),\ 0 \leqq t \leqq 1,$$

all involve quantities with non-vanishing imaginary parts. Hence

$$\begin{aligned}[\arg(\lambda_j\frac{\partial x^1}{\partial \xi^1} + \frac{\partial x^1}{\partial \xi^2})]_{\dot{D}_\rho} &= [\arg(\operatorname{sgn}(\operatorname{Im}\lambda_j)\cdot i\frac{\partial x^1}{\partial \xi^1} + \frac{\partial x^1}{\partial \xi^2})]_{\dot{D}_\rho} \\ &= [\arg(\frac{\partial x^1}{\partial \xi^1} - i\cdot\operatorname{sgn}(\operatorname{Im}\lambda_j)\frac{\partial x^1}{\partial \xi^2})]_{\dot{D}_\rho}.\end{aligned}$$

However, $\frac{\partial x^1}{\partial \xi^1} - i\frac{\partial x^1}{\partial \xi^2} = \frac{dz}{d\zeta}$. By applying (5.5), we obtain

$$\begin{aligned}&\tfrac{1}{2\pi}[\arg\det(A\frac{\partial x^1}{\partial \xi^1} + 1\cdot\frac{\partial x^1}{\partial \xi^2})]_{\dot{D}_\rho} \\ &\qquad = \tfrac{1}{2\pi}\sum_{j=1}^{n}[\arg\frac{\partial x^1}{\partial \xi^1} - i\cdot\operatorname{sgn}(\operatorname{Im}\lambda_j)\frac{\partial x^1}{\partial \xi^2}]_{\dot{D}_\rho} \\ &\qquad = \tfrac{1}{2\pi}(P-N)[\arg\frac{dz}{d\zeta}]_{\dot{D}_\rho} = 2(P-N).\end{aligned}$$

LEMMA 5.3. *Let T be a non-singular matrix, for which*

$$T^{-1}AT = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix},$$

where all of the eigenvalues of C_1 have positive imaginary part, while all of the eigenvalues of C_2 have negative imaginary part. Let $\phi d\zeta$ be a complex cotangent vector. Then, the expression

$$\left\{ A^1T + iA^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} \right\} \begin{pmatrix} \phi \cdot 1_P & 0 \\ 0 & \bar{\phi} \cdot 1_N \end{pmatrix}$$

is conformally invariant.

Proof. Set $\hat{A}^j = \sum A^k \dfrac{\partial x^j}{\partial \xi^k}$, so that

$$A^j = \sum \hat{A}^k \frac{\partial \xi^j}{\partial x^k}.$$

With $\hat{\phi} = \phi \dfrac{d\zeta}{dz}$,

$$\left\{ \hat{A}^1T + i\hat{A}^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} \right\} \begin{pmatrix} \hat{\phi} \cdot 1_P & 0 \\ 0 & \bar{\hat{\phi}} \cdot 1_N \end{pmatrix}$$

$$= \left\{ \hat{A}^1T + i\hat{A}^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} \right\} \begin{pmatrix} \left(\dfrac{\partial \xi^1}{\partial x^1} - i\dfrac{\partial \xi^1}{\partial x^2}\right) \cdot 1_P & 0 \\ 0 & \left(\dfrac{\partial \xi^1}{\partial x^1} + i\dfrac{\partial \xi^1}{\partial x^2}\right) \cdot 1_N \end{pmatrix} \begin{pmatrix} \phi \cdot 1_P & 0 \\ 0 & \bar{\phi} \cdot 1_N \end{pmatrix}$$

$$= \left\{ \hat{A}^1T \frac{\partial \xi^1}{\partial x^1} + \hat{A}^2T \frac{\partial \xi^1}{\partial x^2} + i\left(\hat{A}^1T \frac{\partial \xi^1}{\partial x^2} + \hat{A}^2T \frac{\partial \xi^1}{\partial x^1} \right) \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} \right\} \begin{pmatrix} \phi \cdot 1_P & 0 \\ 0 & \bar{\phi} \cdot 1_N \end{pmatrix}$$

$$= \left\{ A^1T + iA^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} \right\} \begin{pmatrix} \phi \cdot 1_P & 0 \\ 0 & \bar{\phi} \cdot 1_N \end{pmatrix}.$$

(We have used the Cauchy-Riemann equations to obtain the last expression.)

LEMMA 5.4. *If $\phi d\zeta$ is a regular analytic, non-vanishing differential on $\mathcal{S} - D_\rho$ and the genus of $\mathcal{S}$ is p, then*

$$\tfrac{1}{2\pi} [\arg \phi]_{\dot{D}_\rho} = 2p.$$

Proof. Let $g d\zeta = \dfrac{\partial G}{d\zeta} d\zeta$, where G is Green's function of $\mathcal{S} - D_\rho$ [cf. 18]. $\mathcal{S} - D_\rho$ is a surface of genus p with one boundary component. Its double [18] is of genus $2p$. The differential $g d\zeta$ may be extended to the double by means of the Schwarz reflection principle. On the double it represents a divisor of order $4p - 2$. On the boundary of $\mathcal{S} - D_\rho$ it is regular and does not vanish. Hence on $\mathcal{S} - D_\rho$, $g d\zeta$ represents a divisor of order $2p - 1$. Further,

$$[\arg g\frac{d\zeta}{ds}]_{\dot D_\rho}=0, \tfrac{1}{2\pi}[\arg \frac{d\zeta}{ds}]_{\dot D_\rho}=-1.$$

Thus,

$$\tfrac{1}{2\pi}[\arg \phi\frac{d\zeta}{ds}]_{\dot D_\rho}=\tfrac{1}{2\pi}[\arg \phi/g]_{\dot D_\rho}+\tfrac{1}{2\pi}[\arg g\frac{d\zeta}{ds}]_{\dot D_\rho}$$
$$=\tfrac{1}{2\pi}[\arg \phi/g]_{\dot D_\rho}=2p-1.$$

Therefore, $\tfrac{1}{2\pi}[\arg\phi]_{\dot D_\rho}=2p-1-\tfrac{1}{2\pi}[\arg\frac{d\zeta}{ds}]_{\dot D_\rho}=2p.$

LEMMA 5.5. $\tfrac{1}{2\pi}[\arg\det\hat A^1]_{\dot D_\rho}=\tfrac{1}{2\pi}[\arg\det\hat A^2]_{\dot D_\rho}=(2-2p)(P-N).$

Proof. We have $\hat A^1=\hat A^2\hat A$, hence

$$[\arg\det\hat A^1]_{\dot D_\rho}=[\arg\det\hat A^2]_{\dot D_\rho}+[\arg\det\hat A]_{\dot D_\rho}.$$

According to Lemma 5.1, $[\arg\det\hat A]_{\dot D_\rho}=0$, so that

$$[\arg\det\hat A^1]_{\dot D_\rho}=[\arg\det\hat A_2]_{\dot D_\rho}.$$

Next, we observe that

$$\hat A^1=A^1\frac{\partial x^1}{\partial\xi^1}+A^2\frac{\partial x^1}{\partial\xi^2}=A^2(A\frac{\partial x^1}{\partial\xi^1}+1\cdot\frac{\partial x^1}{\partial\xi^2}).$$

Thus

$$\tfrac{1}{2\pi}[\arg\det\hat A^1]_{\dot D_\rho}=\tfrac{1}{2\pi}[\arg\det A^2]_{\dot D_\rho}$$
$$+\tfrac{1}{2\pi}[\arg\det(A\frac{\partial x^1}{\partial\xi^1}+1\cdot\frac{\partial x^1}{\partial\xi^2})]_{\dot D_\rho}.$$

By Lemma 5.2,

$$\tfrac{1}{2\pi}[\arg\det(A\frac{\partial x^1}{\partial\xi^1}+1\cdot\frac{\partial x^1}{\partial\xi^2})]_{\dot D_\rho}=2(P-N).$$

With T locally defined as in Lemma 5.3, (cf. § 5.B, (5.10))

$$T^{-1}A^1T+iT^{-1}A^2T\begin{pmatrix}1_P & 0\\ 0 & -1_N\end{pmatrix}$$
$$=T^{-1}A^2T(T^{-1}AT+i\begin{pmatrix}1_P & 0\\ 0 & -1_N\end{pmatrix})$$
$$=T^{-1}A^2T\begin{pmatrix}C_1+i\cdot 1_P & 0\\ 0 & C_2-i\cdot 1_N\end{pmatrix}$$

Hence,

$$[\arg\det\{T^{-1}A^1T+iT^{-1}A^2T\begin{pmatrix}1_P & 0\\ 0 & -1_N\end{pmatrix}\}]_{\dot D_\rho}$$
$$=[\arg\det A^2]_{\dot D_\rho}+[\arg\det(C_1+i\cdot 1_P)]_{\dot D_\rho}+[\arg\det(C_2-i\cdot 1_N]_{\dot D_\rho}.$$

However, if $\lambda_1, \cdots, \lambda_P$ represent the eigenvalues of C_1, $\mu_1, \cdots, \mu_N$, the eigenvalues of C_2, we have $\operatorname{Im} \lambda_j > 0$, $\operatorname{Im} \mu_k < 0$, and

$$[\arg \det (C_1 + i \cdot 1_P)]_{\dot{D}_\rho} = \sum_{j=1}^{P} [\arg (\lambda_j + i)]_{\dot{D}_\rho} = 0$$

$$[\arg \det (C_2 - i \cdot 1_N)]_{\dot{D}_\rho} = \sum_{j=1}^{N} [\arg (\mu_k - i)]_{\dot{D}_\rho} = 0.$$

Thus

$$[\arg \det A^2]_{\dot{D}_\rho} = [\arg \det \{T^{-1}A^1T + iT^{-1}A^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix}\}]_{\dot{D}_\rho}.$$

Incidentally,

$$\begin{aligned} &\det\{T^{-1}A^1T + iT^{-1}A^2T\} \\ &\quad = \det(T^{-1}A^2T) \cdot \det \begin{pmatrix} C_1 + i \cdot 1_P & 0 \\ 0 & C_2 - i \cdot 1_N \end{pmatrix} \\ &\quad = (\det A^2) \prod_{j=1}^{P} (\lambda_j + i) \prod_{k=1}^{N} (\mu_k - i) \end{aligned}$$

has a value which does not depend on the normalizing matrix T.

With T locally defined, and $\phi d\zeta$ as in Lemma 5.4, we must have

$$[\arg \det \{T^{-1}A^1T + iT^{-1}A^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix}\} \begin{pmatrix} \phi \cdot 1_P & 0 \\ 0 & \bar{\phi} \cdot 1_N \end{pmatrix}]_{\dot{\Delta}} = 0$$

along the boundary $\dot{\Delta}$ of a small triangle Δ on $\mathcal{S} - D_\rho$. Using the conformal invariance established in Lemma 5.3, and a sufficiently fine triangulation of $\mathcal{S} - D_\rho$ (however, see the discussion of § 5.B after (5.11)), we obtain

$$\begin{aligned} 0 &= [\arg \det \{T^{-1}A^1T + iT^{-1}A^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix}\} \begin{pmatrix} \phi \cdot 1_P & 0 \\ 0 & \bar{\phi} \cdot 1_N \end{pmatrix}]_{\dot{D}_\rho} \\ &= [\arg \det \{T^{-1}A^1T + iT^{-1}A^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix}\}]_{\dot{D}_\rho} + P[\arg \phi]_{\dot{D}_\rho} - N[\arg \phi]_{\dot{D}_\rho}. \end{aligned}$$

By Lemma 5.4, $\frac{1}{2\pi}[\arg \phi]_{\dot{D}_\rho} = 2p$. Hence,

$$\frac{1}{2\pi}[\arg \det \{T^{-1}A^1T + iT^{-1}A^2T \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix}\}]_{\dot{D}_\rho} = \frac{1}{2\pi}[\arg \det A^2]_{\dot{D}_\rho} = -2p(P-N),$$

and

$$\frac{1}{2\pi}[\arg \det \hat{A}^1]_{\dot{D}_\rho} = (2-2p)(P-N).$$

As a consequence of Lemma 5.5, we immediately have

THEOREM 5.1. *For a given first order elliptic operator defined on a closed orientable surface of genus $p \neq 1$, $P = N$.*

On the torus, however, one can actually construct first order elliptic operators where P and N are arbitrary.

B. *The principal fiber bundle associated with an operator.*

With every given elliptic operator of first order on the surface $\mathcal{S}$, we shall associate a principal fibre bundle with structural group $GL(P,C)$. In order to define this bundle, we consider the matrix $A=(A^2)^{-1}A^1$. At each point of $\mathcal{S}$, we now consider all non-singular matrices T, for which

$$T^{-1}AT=\begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}, \tag{5.7}$$

as in Lemma 5.3. If we write

$$T=(T_1,T_2), \tag{5.8}$$

where T_1 is an $n\times P$ matrix, T_2 an $n\times N$ matrix, we see that the columns of T_1 form a basis for the space spanned by the eigenvectors of A corresponding to eigenvalues with positive imaginary part, while the columns of T_2 form a basis for the space spanned by the eigenvectors corresponding to eigenvalues with negative imaginary part. Observe that under a change of local coördinates,

$$\hat{A}=(-A\frac{\partial x^1}{\partial\xi^2}+1\cdot\frac{\partial x^1}{\partial\xi^1})^{-1}(A\frac{\partial x^1}{\partial\xi^1}+1\cdot\frac{\partial x^1}{\partial\xi^2}), \tag{5.9}$$

while

$$T^{-1}\hat{A}T = \begin{pmatrix} (-C_1\frac{\partial x^1}{\partial\xi^2}+1_P\cdot\frac{\partial x^1}{\partial\xi^1})^{-1}(C_1\frac{\partial x^1}{\partial\xi^1}+1_P\cdot\frac{\partial x^1}{\partial\xi^2}) & 0 \\ 0 & (-C_2\frac{\partial x^1}{\partial\xi^2}+1_N\cdot\frac{\partial x^1}{\partial\xi^1})^{-1}(C_2\frac{\partial x^1}{\partial\xi^1}+1_N\frac{\partial x^1}{\partial\xi^2}) \end{pmatrix}$$

The matrix $(-C_1\frac{\partial x^1}{\partial\xi^2}+1_P\cdot\frac{\partial x^1}{\partial\xi^1})^{-1}(C_1\frac{\partial x^1}{\partial\xi^1}+1_P\cdot\frac{\partial x^1}{\partial\xi^2})$ has the eigenvalues

$$(-\lambda_j\frac{\partial x^1}{\partial\xi^2}+\frac{\partial x^1}{\partial\xi^1})^{-1}(\lambda_j\frac{\partial x^1}{\partial\xi^1}+\frac{\partial x^1}{\partial\xi^2}),\quad j=1,\cdots,P.$$

(We use the notation of Lemma 5.5.) Furthermore,

$$\frac{\lambda_j\partial x^1/\partial\xi^1+\partial x^1/\partial\xi^2}{-\lambda_j\partial x^1/\partial\xi^2+\partial x^1/\partial\xi^1}=\frac{(\lambda_j\partial x^1/\partial\xi^1+\partial x^1/\partial\xi^2)(-\bar{\lambda}_j\partial x^1/\partial\xi^2+\partial x^1/\partial\xi^1)}{|-\lambda_j\partial x^1/\partial\xi^2+\partial x^1/\partial\xi^1|^2}$$

But the imaginary part of the numerator in the last expression is

$$(\operatorname{Im}\lambda_j)\{\left(\frac{\partial x^1}{\partial\xi^1}\right)^2+\left(\frac{\partial x^1}{\partial\xi^2}\right)^2\}>0.$$

In the same way, we see that the eigenvalues

$$\frac{\mu_k \partial x^1/\partial \xi^1 + \partial x^1/\partial \xi^2}{-\mu_k \partial x^1/\partial \xi^2 + \partial x^1/\partial \xi^1}, \quad k = 1, \cdots, N,$$

of the corresponding $N \times N$ matrix have negative imaginary part. Therefore $T^{-1}\hat{A}T$ also has the required normal form. Consequently, the vector spaces spanned by the columns of T_1, T_2 are indeed defined quite independently of the local coördinate system. The columns of T_1 can, for example, be obtained by considering the projection

$$\chi = \frac{1}{2\pi i}\oint (z \cdot 1 - A)^{-1} dz, \tag{5.10}$$

where the contour of integration surrounds all eigenvalues in the upper half-plane and no others. The columns of χ span the space spanned by the eigenvectors corresponding to eigenvalues with positive imaginary part. Formula (5.10) shows further that if a certain set of columns of χ forms a basis for this space at a point of $\mathcal{S}$, then this same set of columns yields a basis for the corresponding spaces at neighboring points. Thus, locally, we may choose a smooth set of bases.

Assume now that we have an open covering $\{U_\alpha\}$ of $\mathcal{S}$, such that in U_α, the $n \times P$ matrix function $T_{1,\alpha}$ is of class C^∞. If $U_\alpha \cap U_\beta \neq \emptyset$, then we have the two functions $T_{1,\alpha}$ and $T_{1,\beta}$ defined in the intersection. Of course, $T_{1,\alpha}g_{\alpha\beta} = T_{1,\beta}$, where $g_{\alpha\beta}$ is a uniquely determined non-singular $P \times P$ matrix. The matrix $g_{\alpha\beta}$ is again of class C^∞ in $U_\alpha \cap U_\beta$. The functions $g_{\alpha\beta}$ serve as transition functions for a fiber bundle, since for $U_\alpha \cap U_\beta \cap U_\gamma \neq \emptyset$,

$$T_{1,\alpha}g_{\alpha\gamma} = T_{1,\beta}g_{\beta\gamma} = T_{1,\alpha}g_{\alpha\beta}g_{\beta\gamma},$$

so that $g_{\alpha\gamma} = g_{\alpha\beta}g_{\beta\gamma}$.

Of course, at a point of U_α, an arbitrary T_1 may be obtained from $T_{1,\alpha}$ by means of a uniquely defined element $g_\alpha \in GL(P, C)$, so that $T_{1,\alpha}g_\alpha = T_1$. If the point in question lies in $U_\alpha \cap U_\beta$, then also $T_{1,\beta}g_\beta = T_1$. But $T_{1,\beta}g_\beta = T_{1,\alpha}g_{\alpha\beta}g_\beta$, so that $g_\alpha = g_{\alpha\beta}g_\beta$. We therefore use the group $GL(P, C)$ as typical fiber, where $GL(P, C)$ acts on itself from the *left*. By means of the standard construction [22, p. 14], we obtain a principal fiber bundle over $\mathcal{S}$, with structural group $GL(P, C)$. This will be the principal fiber bundle which we associate with the first order differential operator.

We now make an observation. We could have constructed, in the same way, a principal fiber bundle over $\mathcal{S}$ with structural group $GL(N, C)$. We shall show that these two principal fiber bundles are related. In fact, the first Chern class of the second bundle is simply the negative of the first Chern class of the first bundle.

In order to arrive at this result, we begin with the following analysis. Let us triangulate the surface $\mathcal{S}$ by means of triangles $\{\Delta_\alpha\}$. In each Δ_α, we choose smooth matrices $T_{1,\alpha}$ and $T_{2,\alpha}$. We now give each of the 1-simplices of our triangulation an arbitrary, but fixed orientation. Each 1-simplex is in incidence with two 2-simplices. Call the *left* 2-simplex Δ_L and the right 2-simplex Δ_R. We now compute the number

$$\omega = \tfrac{1}{2\pi} \sum [\arg \det {}^t\bar{T}_{1,L} T_{1,R}], \tag{5.11}$$

where the summation is extended over all 1-simplices. This number is called the *characteristic* of the first order elliptic operator [23].

We observe that a much simpler expression for ω can be obtained if we take advantage of the surface $\mathcal{S} - D_\rho$. If we restrict our first bundle to $\mathcal{S} - D_\rho$, then *it has a cross section* and thus is trivial. This follows from the fact that the second cohomology group of $\mathcal{S} - D_\rho$ is trivial, so that the obstruction cocycle [22] of the restriction of our bundle to $\mathcal{S} - D_\rho$ vanishes. Consequently, we have a cross section which has the local representation $\{g_\alpha\}$, with $g_\beta = g_{\beta\alpha} g_\alpha$. Thus, if we set $W_\alpha = T_{1,\alpha} g_\alpha$, then

$$W_\beta = T_{1,\beta} g_\beta = T_{1,\beta} g_{\beta\alpha} g_\alpha = T_{1,\alpha} g_\alpha = W_\alpha,$$

i.e., we may define $T_1^- = W_\alpha$ in U_α, so that T_1^- is continuously defined on $\mathcal{S} - D_\rho$. Similarly, we may define T_1^+ over $\bar{D}_\rho$. (In fact, $\bar{D}_\rho$ is contractible [22].) On $\dot{D}_\rho$, $T_1^- = T_1^+ g$, and

$$\omega = \tfrac{1}{2\pi} [\arg \det {}^t\bar{T}_1^+ T_1^-]_{\dot{D}_\rho}. \tag{5.12}$$

On the other hand,

$${}^t\bar{T}_1^+ T_1^- = {}^t\bar{T}_1^+ T_1^+ g.$$

The matrix ${}^t\bar{T}_1^+ T_1^+$ is Hermitian and non-singular. (In fact, this is just the Gram matrix of the column vectors of T_1^+.) Thus $\det {}^t\bar{T}_1^+ T_1$ is always real and non-zero. Consequently

$$\omega = \tfrac{1}{2\pi} [\arg \det g]_{\dot{D}_\rho}. \tag{5.13}$$

ω is then just the value of the obstruction cocycle of the first bundle on the fundamental cycle of $\mathcal{S}$. The cohomology class represented by the obstruction cocycle is the first Chern class of the bundle [14, 22].

A similar construction for the second bundle yields $T_2^- = T_2^+ h$ on $\dot{D}_\rho$. Hence

$$T^- = (T_1^-, T_2^-) = T^+ \begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix} . \tag{5.14}$$

Now

$$(5.15)\qquad [\arg\det {}^t\bar{T}^+T^-]_{\dot{D}_\rho} = [\arg\det T^-]_{\dot{D}_\rho} - [\arg\det T^+]_{\dot{D}_\rho} = 0 - 0 = 0.$$

(The result for T^- is obtained from a triangulation of $\mathcal{S} - D_\rho$.) But

$$(5.16)\qquad {}^t\bar{T}^+T^- = {}^t\bar{T}^+T^+\begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix},$$

so that

$$(5.17)\qquad [\arg\det {}^t\bar{T}^+T^-]_{\dot{D}_\rho} = [\arg\det {}^t\bar{T}^+T^+]_{\dot{D}_\rho} + [\arg\det g]_{\dot{D}_\rho} + [\arg\det h]_{\dot{D}_\rho}.$$

On the other hand,

$$[\arg\det {}^t\bar{T}^+T^+]_{\dot{D}_\rho} = [\arg\det T^+]_{\dot{D}_\rho} - [\arg\det T^+]_{\dot{D}_\rho} = 0.$$

(5.15) and (5.17) therefore yield

$$(5.18)\qquad [\arg\det g]_{\dot{D}_\rho} + [\arg\det h]_{\dot{D}_\rho} = 0.$$

This proves our assertion about the Chern classes.

C. *Special operators of arbitrary characteristic.*

In this section we construct special first order elliptic operators of arbitrary characteristic ω. Since we are interested primarily in examples with non-zero characteristic, we assume that $0 < P < n$. We begin with the special matrix

$$(5.19)\qquad i\begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix}.$$

On $\mathcal{S} - D_\rho$, we choose

$$(5.20)\qquad A = i\begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix},\quad T_1^- = \begin{pmatrix} 1_P \\ 0 \end{pmatrix},\quad T_2^- = \begin{pmatrix} 0 \\ 1_N \end{pmatrix},$$

so that T^- is the $n \times n$ unit matrix. On $\dot{D}_\rho$, we choose as boundary values for T^+, the matrix

$$(5.21)\qquad T^+ = \begin{bmatrix} z^{-\omega} & & & & & & & 0 \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & & z^{\omega} & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ 0 & & & & & & & 1 \end{bmatrix},$$

so that

$$T_1{}^+ = \begin{pmatrix} z^{-\omega} & 0 \\ 0 & 1_{P-1} \\ 0 & 0 \end{pmatrix}, \quad T_2{}^+ = \begin{pmatrix} 0 & 0 \\ z^{\omega} & 0 \\ 0 & 1_{N-1} \end{pmatrix}. \tag{5.22}$$

On $\dot{D}_\rho$, $\det T^+ = 1$, so that

$$[\arg\det T^+]_{\dot{D}_\rho} = 0. \tag{5.23}$$

Hence T^+ possesses a continuation into the interior of D_ρ, with $\det T^+ \neq 0$ everywhere. (The argument of Steenrod [22, p. 25] allows one to construct a continuation of class C^∞.) Further, $T_1{}^+ = T_1{}^- g_1$, $T_2{}^+ = T_2{}^- g_2$ on $\dot{D}_\rho$, with

$$g_1 = \begin{pmatrix} z^{-\omega} & 0 \\ 0 & 1_{P-1} \end{pmatrix}, \quad g_2 = \begin{pmatrix} z^{\omega} & 0 \\ 0 & 1_{N-1} \end{pmatrix}. \tag{5.24}$$

We now define

$$\hat{A} = iT^+ \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} (T^+)^{-1} \tag{5.25}$$

on D_ρ. This gives a continuation of A onto all of $\mathcal{S}$. On $\mathcal{S} - D_\rho$, the operator

$$A\frac{\partial u}{\partial \xi^1} + 1\cdot\frac{\partial u}{\partial \xi^2} = i\left\{ \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} \frac{\partial u}{\partial \xi^1} - i\cdot 1\frac{\partial u}{\partial \xi^2} \right\} \tag{5.26}$$

thus yields a system of N Cauchy-Riemann operators and P conjugate Cauchy-Riemann operators. We now consider a non-vanishing differential $\phi d\zeta$ on $\mathcal{S} - D_\rho$, and we form the elliptic system

$$\begin{pmatrix} \phi^{-1}\cdot 1_P & 0 \\ 0 & \bar{\phi}^{-1}\cdot 1_N \end{pmatrix} \{A\frac{\partial u}{\partial \xi^1} + 1\cdot\frac{\partial u}{\partial \xi^2}\} = A^1\frac{\partial u}{\partial \xi^1} + A^2\frac{\partial u}{\partial \xi^2} \tag{5.27}$$

relative to local coördinate systems on $\mathcal{S} - D_\rho$. We observe that on $\dot{D}_\rho$, we have, according to Lemma 5.5,

$$\tfrac{1}{2\pi}[\arg\det \hat{A}^j]_{\dot{D}'} = (2-2p)(P-N), \quad j = 1, 2.$$

If $(2-2p)(P-N) = 0$, the matrix $\hat{A}^2$ may be continued as a non-singular matrix into D_ρ. We then set $\hat{A}^1 = \hat{A}^2\hat{A}$ in D_ρ, where $\hat{A}$ is defined through (5.25). (Note that $\hat{A} = A$ on $D - D_\rho$ in our special situation.) Thus we have constructed an elliptic system on $\mathcal{S}$, for which

$$[\arg\det {}^t\bar{T}_1{}^+ T_1{}^-]_{\dot{D}'} = [\arg\det(g_1)^{-1}]_{\dot{D}_\rho} = 2\pi\omega.$$

The characteristic of the system, therefore, is ω.

D. *The index of a special operator.*

We now consider a *special* equation of characteristic ω,

$$Lu \equiv A^1 \frac{\partial u}{\partial x^1} + A^2 \frac{\partial u}{\partial x^2} = f. \tag{5.28}$$

After changing the dependent variables u, so that

$$u = \begin{cases} T^- w^- \text{ in } \mathcal{S} - D_\rho \\ T^+ w^+ \text{ in } D_\rho, \end{cases} \tag{5.29}$$

we have

$$\mathcal{L} w \equiv i \begin{pmatrix} 1_P & 0 \\ 0 & -1_N \end{pmatrix} \frac{\partial w}{\partial x^1} + \frac{\partial w}{\partial x^2} + \text{lower order terms} \\ = T^{-1}(A^2)^{-1} f, \tag{5.30}$$

with

$$T^+ w^+ = T^- w^- \text{ on } \dot{D}_\rho. \tag{5.31}$$

With our special choices for $T^\pm$, these boundary conditions are simply

$$w_1^+ = z^\omega w_1^-, \quad w_{P+1}^+ = z^{-\omega} w_{P+1}^-, \; w_j^+ = w_j^-, \quad j \neq 1, P+1. \tag{5.32}$$

The range of the operator $\mathcal{L}$ in (5.30) consists of $n \times 1$ vectors v, whose components are coefficients of forms of type $(1,0)$ or $(0,1)$. This follows from the transformation law for $A^2 T$ in our special situation. In fact,

$$v_1 dz, \cdots, v_P dz, \; v_{P+1} d\bar{z}, \cdots, v_n d\bar{z} \text{ are conformally invariant.}$$

We now consider the Hilbert space $H_{1,n}(\mathcal{S})$ which is the completion of the piecewise smooth C^∞, $n \times 1$ vectors w on $\mathcal{S}$ which satisfy (5.31), relative to the norm $\| \; \|_{1,n}$. The operator $\mathcal{L}$ can then be viewed as a bounded operator of the space $H_{1,n}(\mathcal{S})$ into the Hilbert space $\tilde{H}_{0,n}(\mathcal{S})$ of square integrable vectors v described above [cf. 16]. We denote by $\mathcal{L}_0$ that operator which arises from $\mathcal{L}$ by discarding the lower order terms in (5.30). $\mathcal{L}_0$ is then also a bounded linear transformation from $H_{1,n}(\mathcal{S})$ into $\tilde{H}_{0,n}(\mathcal{S})$, which differs from $\mathcal{L}$ only by a compact linear transformation. Further, a coercive inequality of the form

$$\| w \|_{1,n} \leqq C(\| \mathcal{L}_0 w \|_{0,n} + \| w \|_{0,n}) \tag{5.33}$$

may be derived for all $w \in H_{1,n}(\mathcal{S})$ [16]. Actually, $\mathcal{L}_0$ is a Φ-operator whose index may be computed by means of the Riemann-Roch theorem [16]. (This follows from the nature of the boundary conditions (5.32), which allow one to compute the index by discussing analytic functions on $\mathcal{S}$, which are multiples of a certain divisor.) This index is given by

$$-2\omega + n(1-p). \tag{5.34}$$

Of course, by Theorem 3.1, the index of $\mathcal{L}$ is also given by (5.34).

Next, we observe that (5.29) establishes a 1-1 correspondence between piecewise smooth C^∞ functions u which satisfy

$$u^+ = u^- \text{ on } \dot{D}_\rho, \tag{5.35}$$

and piecewise smooth C^∞ function w which satisfy (5.31). Because of (5.35), it is easily seen that for any test vector ϕ of class C^∞, with support contained in a coördinate neighborhood of $\mathcal{S}$,

$$\int {}^t u \phi_{x_j} \, dx = -\int ({}^t u)_{x_j} \phi \, dx. \tag{5.36}$$

Because of the identity of strong and weak derivatives [11], it follows that such a function u belongs to $L^2{}_{1,n}(\mathcal{S})$. On the other hand, every vector function u which is of class C^∞ on all of $\mathcal{S}$, certainly satisfies (5.35). Further, these C^∞ function are dense in $L^2{}_{1,n}(\mathcal{S})$. Consequently $L^2{}_{1,n}(\mathcal{S})$ is the completion with respect to the norm $\| \quad \|_{1,n}$ of the space of piecewise smooth functions u which satisfy (5.35). In view of (5.29), one can derive inequalities

$$C^{-1} \| w \|_{1,n} \leqq \| u \|_{1,n} \leqq C \| w \|_{1,n} \tag{5.37}$$

for corresponding piecewise smooth functions u and w. (The constant C does not depend on u and w.) From this discussion it follows that the spaces $H_{1,n}(\mathcal{S})$ and $L^2{}_{1,n}(\mathcal{S})$ are topologically isomorphic. Further, the correspondence

$$v \to A^2 T v \tag{5.38}$$

yields a topological isomorphism between $\tilde{H}_{0,n}(\mathcal{S})$ and $L^2{}_{,n}(\mathcal{S})$. This isomorphism maps the range of the operator $\mathcal{L}$ onto the range of the operator L. (We remark that L is originally defined only on the functions of class C^∞. These are dense in $L^2{}_{1,n}(\mathcal{S})$. As a bounded operator, L has a *unique* extension to all of $L^2{}_{1,n}(\mathcal{S})$. On the other hand, the original definition for L also makes sense on the piecewise smooth functions u satisfying (5.35). However, L is a bounded operator on these functions.) Hence, we have

$$\tilde{H}_{0,n}(\mathcal{S})/\text{range}\, \mathcal{L} \cong L^2{}_{,n}(\mathcal{S})/\text{range}\, L. \tag{5.39}$$

Further $\ker L \cong \ker \mathcal{L}$. Consequently, $\operatorname{ind} L = \operatorname{ind} \mathcal{L}$. We have therefore established

THEOREM 5.2. *The index of a special $n \times n$ elliptic operator of first order of characteristic ω, on a surface $\mathcal{S}$ of genus p, is equal to*

$$-2\omega + n(1-p).$$

Remark. It is possible to establish Theorem 5.2 also for general first order elliptic operators of characteristic ω. This result, together with Theorem 5.1 shows that the index of an elliptic system of first order is always an *even* integer.

6. The degree of a characteristic matrix. In § 4, we associated with each continuous mapping

$$\mathfrak{A}: \mathfrak{B} \to GL(n, C)$$

an integer $\kappa(\mathfrak{A})$ defined by the index of a related singular integral operator. In this section, we shall define a second integer $l(\mathfrak{A})$, called the degree of the mapping $\mathfrak{A}$. In order to do this, we consider the polar decomposition of $\mathfrak{A}$,

$$\mathfrak{A} = ({}^t\bar{\mathfrak{A}}\mathfrak{A})^{\frac{1}{2}} U, \tag{6.1}$$

where U is a unitary matrix. Set $B = {}^t\bar{\mathfrak{A}}\mathfrak{A}$, and consider

$$B^{\frac{1}{2}} = \frac{1}{2\pi i}\oint z^{\frac{1}{2}}(z \cdot 1 - B)^{-1} dz, \tag{6.2}$$

where we integrate around the spectrum of B, while $z = 0$ remains outside the contour. We take the principal branch of $z^{\frac{1}{2}}$. Thus we see that $B^{\frac{1}{2}}$ depends smoothly on B. Hence U is a continuous function on the cotangent sphere bundle $\mathfrak{B}$ of $\mathcal{S}$. Observe that the dimension of $\mathfrak{B}$ is three and that $\mathfrak{B}$ is an orientable manifold because $\mathcal{S}$ is an orientable surface. We have the mapping

$$U: \mathfrak{B} \to U(n), \tag{6.3}$$

which we compose with the projection

$$\pi: U(n) \to U(n)/U(1). \tag{6.4}$$

(We use the convention that for each $k < n$, $U(k)$ is that subgroup of $U(n)$ which is given by the imbedding

$$\begin{pmatrix} 1_{n-k} & 0 \\ 0 & U(k) \end{pmatrix}.)$$

We observe that the third homology group of $U(n)/U(1)$ is infinite cyclic [22, p. 134].

Definition 6.1. The *degree* $l(\mathfrak{A})$ of the mapping $\mathfrak{A}$ is equal to the degree of the mapping $\pi \circ U = f$, i.e. if χ_3 is the generator of $H_3(\mathfrak{B}; Z)$ and $\hat{\chi}_3$ is the generator of $H_3(U(n)/U(1); Z)$, then the degree is the integer j, where $f_*\chi_3 = j\hat{\chi}_3$.

We now wish to compute the degrees of the characteristic forms of some of our special elliptic systems of § 5. C. For this, we restrict ourselves to the case $n=2$. Then $\mathfrak{A}$ is of the form

$$\mathfrak{A}=(a_{ij}),\quad i,j=1,2. \tag{6.5}$$

Suppose that

$$U=\begin{pmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{pmatrix} \tag{6.6}$$

is the unitary matrix which enters the polar decomposition of $\mathfrak{A}$. An element of $U(1)$ has the form

$$\begin{pmatrix}1 & 0\\ 0 & \gamma\end{pmatrix},\quad |\gamma|=1, \tag{6.7}$$

so that

$$\begin{pmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & \gamma\end{pmatrix}=\begin{pmatrix}\alpha_{11} & \alpha_{12}\gamma\\ \alpha_{21} & \alpha_{22}\gamma\end{pmatrix}. \tag{6.8}$$

Hence, the vector

$$\begin{pmatrix}\alpha_{11}\\ \alpha_{21}\end{pmatrix} \tag{6.9}$$

determines the coset of $U(2)/U(1)$ to which U belongs. Of course,

$$|\alpha_{11}|^2+|\alpha_{21}|^2=1. \tag{6.10}$$

Thus, if we set $\alpha_{11}=x^1+ix^2$, $\alpha_{21}=x^3+ix^4$, then the point (x^1,x^2,x^3,x^4) is a point in R^4 which lies on the unit sphere S^3 of that space. Furthermore, it is clear that

$$\hat{U}=U\cdot\begin{pmatrix}1 & 0\\ 0 & (\det U)^{-1}\end{pmatrix} \tag{6.11}$$

gives a mapping of $\mathfrak{B}$ into the special unitary group $SU(2)$. We can represent each coset of $U(2)/U(1)$ uniquely in terms of an element of $SU(2)$. In fact, if the vector (6.9) determines an element of $U(2)/U(1)$, then the corresponding element of $SU(2)$ is given by

$$\begin{pmatrix}\alpha_{11} & -\bar{\alpha}_{21}\\ \alpha_{21} & \bar{\alpha}_{11}\end{pmatrix}. \tag{6.12}$$

We see then that the sphere S^3 may be regarded as the group $SU(2)$.

If we consider the mapping $\mathfrak{A}$ of (6.5), then we can define another mapping $g\colon \mathfrak{B}\to S^3$ as follows: Each point of $\mathfrak{B}$ is mapped into that point $(x^1,x^2,x^3,x^4)\in R^4$, for which

$$\frac{a_{11}}{a}=x^1+ix^2,\ \frac{a_{21}}{a}=x^3+ix^4,\quad a=(|a_{11}|^2+|a_{21}|^2)^{\frac{1}{2}}. \tag{6.13}$$

We claim that the two mappings $f = \pi \circ U$ and g are homotopic. Then $f_* = g_*$ and the degrees of the two mappings are equal. In order to see this, we set

$$(6.14) \qquad B_\lambda = \begin{pmatrix} b_{11}(\lambda) & b_{12}(\lambda) \\ b_{21}(\lambda) & b_{22}(\lambda) \end{pmatrix}$$
$$= \{(1-\lambda)^t\bar{\mathfrak{a}}\mathfrak{a} + \lambda\cdot 1\}^{\frac{1}{2}} \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}, 0 \leqq \lambda \leqq 1.$$

Since the right side of (6.14) is the product of two non-singular matrices,

$$|b_{11}(\lambda)|^2 + |b_{21}(\lambda)|^2 \neq 0,$$

and we can define a mapping $h_\lambda\colon \mathfrak{B} \to S^3$ by applying the scheme (6.13) to the matrix B_λ. h_λ gives the desired homotopy.

We have now associated a mapping $g\colon \mathfrak{B} \to S^3$ with every mapping $\mathfrak{a}\colon \mathfrak{B} \to GL(2,C)$ through (6.13). The degree of the mapping g is equal to $l(\mathfrak{a})$, by the homotopy argument above. Thus we have associated a degree with the first column of $\mathfrak{a}$. In a similar way, we can associate a degree with each row or column of $\mathfrak{a}$. We shall now study the relations between these degrees. We begin with

LEMMA 6.1 [15, p. 69]. *Let X be a triangulable space and $f, g\colon X \to S^3$ two maps of X into S^3. Let $\psi\colon S^3 \to S^2$ be the Hopf map. Then, if the mappings ψf and ψg are homotopic, so are the mappings f and g.*

(The Hopf map is just the fibering

$$U(2)/U(1) \to U(2)/U(1) \times U(1).)$$

LEMMA 6.2. *The column degrees of a mapping $U\colon \mathfrak{B} \to U(2)$ are equal.*

Proof. Beginning with U, we form the map $\hat{U}\colon \mathfrak{B} \to SU(2)$ as in (6.11). Observe that the first columns of U and $\hat{U}$ coincide, while the second column of $\hat{U}$ is the second column of U, multiplied by $\det U^{-1}$. In view of Lemma 6.1, the degrees of the second columns of U and $\hat{U}$ are the same, since they yield homotopic mappings of $\mathfrak{B}$ into S^3. However,

$$\hat{U} = \begin{pmatrix} \alpha_{11} & -\bar{\alpha}_{21} \\ \alpha_{21} & \bar{\alpha}_{11} \end{pmatrix}.$$

Now

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha_{11} & -\bar{\alpha}_{21} \\ \alpha_{21} & \bar{\alpha}_{11} \end{pmatrix} = \begin{pmatrix} -\alpha_{21} & -\bar{\alpha}_{11} \\ \alpha_{11} & -\bar{\alpha}_{21} \end{pmatrix},$$

and the map $\mathfrak{B} \to SU(2)$ given by

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

is homotopically equivalent to the constant map 1. Hence the column degrees of

$$\begin{pmatrix} -\bar{\alpha}_{21} \\ \bar{\alpha}_{11} \end{pmatrix} \text{and} \begin{pmatrix} -\bar{\alpha}_{11} \\ -\bar{\alpha}_{21} \end{pmatrix}$$

coincide. From Lemma 6.1, we see that

$$\begin{pmatrix} -\bar{\alpha}_{11} \\ -\bar{\alpha}_{21} \end{pmatrix} \text{and} \begin{pmatrix} \bar{\alpha}_{11} \\ \bar{\alpha}_{21} \end{pmatrix}$$

are homotopically equivalent and thus have the same degree. Finally, the mapping $x_j \to x_j$, $j \neq j_0$, $x_{j_0} \to -x_{j_0}$, $\sum_j x_j^2 = 1$, of S^3 into itself is of degree -1. Hence the degrees of the mappings

$$\begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix}, \begin{pmatrix} \bar{\alpha}_{11} \\ \bar{\alpha}_{21} \end{pmatrix},$$

are equal.

LEMMA 6.3. *For a mapping $\hat{U}: \mathcal{B} \to SU(2)$,*

$$\text{row degree } \hat{U} = - \text{column degree } \hat{U}.$$

Proof. The matrix $\hat{U}$ is of the form

$$\hat{U} = \begin{pmatrix} \alpha_{11} & -\bar{\alpha}_{21} \\ \alpha_{21} & \bar{\alpha}_{11} \end{pmatrix}.$$

The mapping $(z_1, z_2) \to (z_1, -\bar{z}_2)$ of S^3 into itself is of degree -1.

We now turn to the third cohomotopy group, $\pi^3(\mathcal{B})$ [15]. There exists a homomorphism

$$\pi^3(\mathcal{B}) \to H^3(\mathcal{B}; Z)$$

which is defined as follows. Let $\alpha \in \pi^3(\mathcal{B})$ and pick a map $\phi: \mathcal{B} \to S^3$ which represents α. Let ψ_3 be the generator of $H^3(\mathcal{B}; Z)$, $\hat{\psi}_3$, the generator of $H^3(S^3; Z)$. The image of α is defined as the element $\phi^*(\hat{\psi}_3) \in H^3(\mathcal{B}; Z)$. But $\phi^*(\hat{\psi}_3) = j \cdot \psi_3$. The integer j is just the *degree* of the map ϕ. In fact,

$$\langle \psi_3, \chi_3 \rangle = 1, \qquad \langle \hat{\psi}_3, \hat{\chi}_3 \rangle = 1$$

and

$$\langle \phi^*(\hat{\psi}_3), \chi_3 \rangle = j \langle \psi_3, \chi_3 \rangle = \langle \hat{\psi}_3, \phi_* \chi_3 \rangle = d \langle \hat{\psi}_3, \hat{\chi}_3 \rangle,$$

where d is the degree of ϕ. Hence $d = j$.

LEMMA 6.4 [15, p. 213]. *Let $\hat{U}, \hat{V}: \mathcal{B} \to SU(2)$ $(= S^3)$ be given maps and let α, β denote the elements of $\pi^3(\mathcal{B})$ which are represented by $\hat{U}$, $\hat{V}$, respectively. Let $\alpha\beta \in \pi^3(\mathcal{B})$ be the element which is represented by $\hat{U}\hat{V}$: $\mathcal{B} \to SU(2)$. Then $\alpha\beta = \alpha + \beta$.*

COROLLARY 6.1. *With $\hat{U}$, $\hat{V}$ as above,*

$$\text{row degree } \hat{U}\hat{V} = \text{row degree } \hat{U} + \text{row degree } \hat{V},$$
$$\text{column degree } \hat{U}\hat{V} = \text{column degree } \hat{U} + \text{column degree } \hat{V}.$$

The set of mappings $\hat{U}\colon \mathfrak{B} \to SU(2)$ certainly forms a group under matrix multiplication. In view of (4.7), κ is a homomorphism of this group into the additive group Z of integers. By Corollary 6.1, l is also such a homomorphism. We note that if $l(\hat{U}) = 0$, then $\kappa(\hat{U}) = 0$, by the discussion of § 4. There exists a $\hat{U}_0\colon \mathfrak{B} \to SU(2)$, with $l(\hat{U}_0) = 1$ [15, p. 53]. Suppose that $\kappa(\hat{U}_0) = \gamma$. Then, for any $\hat{U}$, with $l(\hat{U}) = z$, we have $l(\hat{U}) = l(\hat{U}_0^z) = z$, and

$$l(\hat{U}\hat{U}_0^{-z}) = l(\hat{U}) + l(\hat{U}_0^{-z}) = z - z = 0.$$

Hence $\kappa(\hat{U}\hat{U}_0^{-z}) = \kappa(\hat{U}) - z\kappa(\hat{U}_0) = 0$, and

$$\kappa(\hat{U}) = \kappa(\hat{U}_0)z = \gamma l(\hat{U}).$$

LEMMA 6.5. *If $U\colon \mathfrak{B} \to U(2)$ is a given mapping, then*

$$\text{row degree } U = -\text{column degree } U.$$

Proof. Consider the matrix function

$$\hat{V} = U\begin{pmatrix} 1 & 0 \\ 0 & \det U^{-1} \end{pmatrix} \cdot U^{-1}\begin{pmatrix} 1 & 0 \\ 0 & \det U \end{pmatrix}.$$

$\hat{V}$ is a map $\mathfrak{B} \to SU(2)$. We observe that because of (4.7) and (3.5), $\kappa(\hat{V}) = 0$. Further $\kappa(\hat{V}) = 0 = \gamma l(\hat{V})$. Since we have constructed elliptic operators of non-zero index in § 5, we know that $\gamma \neq 0$. Hence, $l(\hat{V}) = 0$. However,

$$0 = l(\hat{V}) = l\{U\begin{pmatrix} 1 & 0 \\ 0 & \det U^{-1} \end{pmatrix}\} + l\{U^{-1}\begin{pmatrix} 1 & 0 \\ 0 & \det U \end{pmatrix}\} = l(U) + l(U^{-1})$$
$$= \text{column degree } U + \text{column degree } U^{-1}.$$

On the other hand,

$$U^{-1}U = 1 = U^{-1}\begin{pmatrix} 1 & 0 \\ 0 & \det U \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & \det U^{-1} \end{pmatrix} U,$$

so that

$$0 = l(1) = l\{U^{-1}\begin{pmatrix} 1 & 0 \\ 0 & \det U \end{pmatrix}\} + l\{\begin{pmatrix} 1 & 0 \\ 0 & \det U^{-1} \end{pmatrix} U\}$$

$$= \text{column degree } U^{-1} - \text{row degree } U.$$

Hence, column degree $U = -$ row degree U.

THEOREM 6.1. *The column degree l is a homomorphism on the group of mappings $\mathfrak{B} \to U(2)$ into the additive group of integers.*

Proof. Consider two mappings $U, V \colon \mathfrak{B} \to U(2)$. If $\phi \colon \mathfrak{B} \to U(1)$ is a given mapping, then by Lemma 6.5,

$$\text{column degree} \begin{pmatrix} 1 & 0 \\ 0 & \phi \end{pmatrix} U = -\text{row degree} \begin{pmatrix} 1 & 0 \\ \phi & 0 \end{pmatrix} U$$

$$= -\text{row degree } U = \text{column degree } U.$$

Hence, also

$$\text{column degree } UV = \text{column degree} \begin{pmatrix} 1 & 0 \\ 0 & \det U^{-1} \end{pmatrix} UV$$

$$= \text{column degree} \begin{pmatrix} 1 & 0 \\ 0 & \det U^{-1} \end{pmatrix} UV \begin{pmatrix} 1 & 0 \\ 0 & \det V^{-1} \end{pmatrix}$$

$$= \text{column degree} \begin{pmatrix} 1 & 0 \\ 0 & \det U^{-1} \end{pmatrix} U$$

$$+ \text{column degree } V \begin{pmatrix} 1 & 0 \\ 0 & \det V^{-1} \end{pmatrix}, \quad (\text{Cor. } 6.1)$$

$$= \text{column degree } U + \text{column degree } V.$$

COROLLARY 6.2. *The column degree l is a homomorphism of the group of mappings $\mathfrak{B} \to GL(2, C)$ into the additive group of integers.*

Proof. Let $\mathfrak{A}_1, \mathfrak{A}_2 \colon \mathfrak{B} \to GL(2, C)$ be given mappings and set $\mathfrak{A} = \mathfrak{A}_1\mathfrak{A}_2$. We have

$$\mathfrak{A}_1 = ({}^t\bar{\mathfrak{A}}_1\mathfrak{A}_1)^{\frac{1}{2}}U_1, \qquad \mathfrak{A}_2 = ({}^t\bar{\mathfrak{A}}_2\mathfrak{A}_2)^{\frac{1}{2}}U_2, \qquad \mathfrak{A} = ({}^t\bar{\mathfrak{A}}\mathfrak{A})^{\frac{1}{2}}U.$$

We can define l through the degrees of the mappings discussed in (6.13), (6.14). But $\mathfrak{A}_j \sim U_j$, $j = 1, 2$, as in (6.14), so that

$$\mathfrak{A} = \mathfrak{A}_1\mathfrak{A}_2 \sim U_1U_2.$$

Hence,

$$l(\mathfrak{A}) = l(U_1U_2) = l(U_1) + l(U_2) = l(\mathfrak{A}_1) + l(\mathfrak{A}_2).$$

Finally, we turn from the discussion of maps $\mathfrak{B} \to U(2)/U(1)$, to a discussion of maps $\mathfrak{B} \to U(n)/U(1)$. Let us consider the commutative bundle diagram

$$
(6.15)\qquad
\begin{tikzcd}
U(n) \arrow[rr, "\pi"] \arrow[dr, "\pi_2"'] & & U(n)/U(1) \arrow[dl, "\pi_1"] \\
& U(n)/U(2) &
\end{tikzcd}
$$

The fiber bundle $\pi_1: U(n)/U(1) \to U(n)/U(2)$ is associated with the principal bundle $\pi_2: U(n) \to U(n)/U(2)$ and has $U(2)/U(1) = S^3$ as typical fiber. If i: $S^3 \to U(n)/U(1)$ is the inclusion map, then it is known [22, p. 134] that

$$i_*: H_3(S^3; Z) \to H_3(U(n)/U(1); Z)$$

is an isomorphism. We now consider the following mappings:

$$
\begin{tikzcd}
\mathcal{B} \arrow[r, "U"] & U(2) \arrow[r, "i_1"] \arrow[d, "\pi_3"] & U(n) \arrow[d, "\pi"] \\
& U(2)/U(1) \arrow[r, "i"] & U(n)/U(1).
\end{tikzcd}
$$

Here i_1 is the imbedding discussed after (6.4). The mapping i maps $U(2)/U(1)$ onto the fiber of π_1: $U(n)/U(1) \to U(n)/U(2)$ which lies above the image of $\pi_2 \circ i_1$. Consequently, $i \circ \pi_3 = \pi \circ i_1$.

Let $\tilde{\chi}_3$ be the generator of $H_3(U(2)/U(1); Z)$. Then, with the notation of Definition 6.1, $\hat{\chi}_3 = i_* \tilde{\chi}_3$. By the definition of l, we have

$$(\pi_3 \circ U)_* \chi_3 = l(U) \tilde{\chi}_3.$$

On the other hand,

$$i_* \pi_{3*} U_* \chi_3 = \pi_* i_{1*} U_* \chi_3 = l(U) \hat{\chi}_3.$$

Consequently, we have the result

$$(6.16)\qquad l(i_1 U) = l(U).$$

We now must discuss the homotopy classification of the maps $\mathcal{B} \to U(n)/U(1)$. Since the homotopy groups $\pi_j(U(n)/U(1))$ are zero for $j \leqq 2$, $U(n)/U(1)$ is 2-connected [15, p. 188]. Assume that $\mathcal{B}$ has been triangulated. Call the resulting complex K. On the 2-skeleton K^2, a map $\phi \mid K^2$, where ϕ: $\mathcal{B} \to U(n)/U(1)$, is homotopic to a constant map, since for the boundary $\dot{\sigma}$ of any 3-cell σ, $\phi \mid \dot{\sigma}$ defines an element of $\pi_2(U(n)/U(1)) = 0$. It follows from the homotopy extension theorem [15, p. 14] that ϕ is homotopic to a map $\tilde{\phi}$: $\mathcal{B} \to U(n)/U(1)$, such that $\tilde{\phi} \mid K^2$ is constant. We now define an element of $H^3(\mathcal{B}; \pi_3(U(n)/U(1)))$ as follows. Consider the

cochain $c(\bar{\phi})$, such that $c(\bar{\phi})(\sigma)=[\bar{\phi}\mid\sigma]$, the homotopy class generated by $\bar{\phi}\mid\sigma$. $c(\bar{\phi})$ is then a cocycle. We denote its cohomology class in $H^3(\mathfrak{B};\pi_3(U(n)/U(1)))$ by $\kappa^3(\phi)=\kappa^3(\bar{\phi})$. According to a known generalization of the Hopf homotopy theorem [15, Cor. 15.3, p. 191], two maps $\phi_1,\phi_2\colon\mathfrak{B}\to U(n)/U(1)$ are homotopic if, and only if $\kappa^3(\phi_1)=\kappa^3(\phi_2)$.

Let α represent the generator of $\pi_3(U(n)/U(1))$. Then we have

$$c(\bar{\phi})(K)=(\sum_\sigma k_\sigma)\alpha,$$

where $k_\sigma\alpha$ corresponds to $[\bar{\phi}\mid\sigma]$. We now apply the Hurewicz isomorphism. $\pi_3(U(n)/U(1))\to H_3(U(n)/U(1);Z)$ which sends $(\sum_\sigma k_\sigma)\alpha$ into $(\sum_\sigma k_\sigma)\hat{\chi}_3$. It becomes clear that the integer $j=\sum_\sigma k_\sigma$ is just the degree of the map ϕ. Consequently, two mappings $\phi_1,\phi_2\colon\mathfrak{B}\to U(n)/U(1)$ are homotopic if and only if they have the same degree. Since we can obtain maps $i\pi_3U$ of arbitrary degree, any map $\phi\colon\mathfrak{B}\to U(n)/U(1)$ is always homotopic to a special map of the form $i\pi_3U$, with $U\colon\mathfrak{B}\to U(2)$.

If therefore $V\colon\mathfrak{B}\to U(n)$ is a given mapping, it follows that $\pi\circ V$ is homotopic to a map of the form $i\pi_3U=\pi i_1U$. By the second covering homotopy theorem [22], the homotopy $\mathfrak{B}\times I\to U(n)/U(1)$ of $\pi\circ V$ can be covered by a homotopy $V_t\colon\mathfrak{B}\times I\to U(n)$. Thus $\pi\circ V_1=\pi\circ i_1U$, i.e. V_1 and i_1U always belong to the same fiber of the bundle $\pi\colon U(n)\to U(n)/U(1)$. Since i_1U is a member of the subgroup

$$\begin{pmatrix}1_{n-2} & 0\\ 0 & U(2)\end{pmatrix},\tag{6.17}$$

for all points of $\mathfrak{B}$, V_1 must also belong to this subgroup for every point of $\mathfrak{B}$. Hence we have shown that every mapping $\mathfrak{B}\to U(n)$ is homotopically equivalent to a mapping of $\mathfrak{B}$ into the subgroup (6.17). This fact allows us to restrict all of our considerations to the special mappings $U\colon\mathfrak{B}\to U(2)$. Formula (6.16) then gives us $l(V)=l(V_1)=l(i_1U)=l(U)$.

7. The identity of the index and the degree. In the discussion which precedes Lemma 6.5, we showed that $\kappa(\hat{U})=\gamma l(\hat{U})$ for every mapping $\hat{U}\colon\mathfrak{B}\to SU(2)$. Since, by § 4 and Theorem 6.1, both κ and l are homomorphisms of the group of mappings $\mathfrak{B}\to U(2)$ into the additive group of integers, then, for any map $U\colon\mathfrak{B}\to U(2)$, we must have

$$\kappa(U)=\kappa(\hat{U}),\quad l(U)=l(\hat{U}),$$

where U and $\hat{U}$ are related as in (6.11). (The result for κ follows from (3.5).) Hence we have

$$\kappa(U) = \gamma l(U) \tag{7.1}$$

for all mappings $U: \mathfrak{B} \to U(2)$. In order to prove that the value of γ is -1, it is sufficient to show this for a single case, with $\kappa(U) \neq 0$.

In § 5. C, we constructed special elliptic operators of first order, of arbitrary characteristic. We now establish

THEOREM 7.1. *The degree of a characteristic matrix of a special elliptic operator of first order with $P = N = 1$, of characteristic zero, is equal to the negative of the Euler characteristic of the surface $\mathcal{S}$ on which the operator is defined.*

Proof. The characteristic matrix of a system of first order has the form

$$A^1\xi_1 + A^2\xi_2$$

for the cotangent vector $\xi_1 dx^1 + \xi_2 dx^2$. If we introduce a new natural base in the fibres of the cotangent bundle in terms of our conformal structure, we have $dz = dx^1 + i dx^2$, $d\bar{z} = dx^1 - i dx^2$, and

$$\xi_1 dx^1 + \xi_2 dx^2 = \tfrac{1}{2}\{(\xi_1 - i\xi_2)\,dz + (\xi_1 + i\xi_2)\,d\bar{z}\} = \phi dz + \bar{\phi}\, d\bar{z},$$

where $\phi = \frac{1}{2}(\xi_1 - i\xi_2)$. Then

$$\begin{aligned} A^1\xi_1 + A^2\xi_2 &= \tfrac{1}{2}(A^1 + iA^2)(\xi_1 - i\xi_2) + \tfrac{1}{2}(A^1 - iA^2)(\xi_1 + i\xi_2) \\ &= (A^1 + iA^2)\phi + (A^1 - iA^2)\bar{\phi}. \end{aligned}$$

However, we have

$$\begin{aligned} &(A^1 + iA^2)\phi + (A^1 - iA^2)\bar{\phi} \\ &= \left\{A^1 + iA^2\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right\}\begin{pmatrix}\phi & 0\\ 0 & \bar{\phi}\end{pmatrix} + \left\{A^1 - iA^2\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right\}\begin{pmatrix}\bar{\phi} & 0\\ 0 & \phi\end{pmatrix}. \end{aligned} \tag{7.2}$$

Further, for a *special system* of characteristic 0, with $P = N = 1$, (we can take $T = 1$ in this case),

$$A = (A^2)^{-1}A^1 = i\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

Hence

$$A^1 - iA^2\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = A^1 - A^2A = A^1 - A^1 = 0.$$

Consequently, the characteristic matrix of our special system of characteristic zero is given by

$$\{A^1 + iA^2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \phi & 0 \\ 0 & \bar{\phi} \end{pmatrix}.$$

Observe now that this matrix function defines a map of the non-zero cotangent vectors of $\mathcal{S}$ into the general linear group $GL(2,C)$. Further, if we combine this mapping with the projection

$$GL(2,C) \to GL(2,C)/GL(1,C),$$

we obtain a mapping of the non-zero contangent vectors into $C_2 - \{0\}$, where the image point is represented by the elements of the first column of the 2×2 matrix, $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$. By dividing by $(|z_1|^2 + |z_2|^2)^{\frac{1}{2}}$, we get a mapping into S^3. It is clear that if t is positive, then

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} z_1 t \\ z_2 t \end{pmatrix}$$

are projected into the same point of S^3. Consequently we get a mapping of the contangent sphere bundle $\mathcal{B}$ into $S^3 = U(2)/U(1)$. We now consider the Hopf fibering

$$S^3 = U(2)/U(1) \to U(2)/U(1) \times U(1) = P_1 = S^2.$$

This is just associated with the canonical complex line bundle of the complex projective space P_1, and it is universal for dimension 2 [22]. The first Chern class of this universal bundle is known to be just the element $\hat{\psi}_2 \in H^2(S^2; Z)$, the generator of this cohomology group. Observe now that our characteristic matrix gives a *bundle map*

$$\begin{array}{ccc} \mathcal{B} & \xrightarrow{\hat{f}} & S^3 \\ \downarrow & & \downarrow \\ \mathcal{S} & \xrightarrow{f} & S^2 \end{array}$$

since each fiber of the bundle $\mathcal{B} \to \mathcal{S}$ is mapped homeomorphically onto a fiber of the universal bundle. (If we denote the elements of

$$A^1 + iA^2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

by (b_{ij}), then the characteristic matrix takes the form

$$\begin{pmatrix} b_{11}\phi & b_{12}\bar{\phi} \\ b_{21}\phi & b_{22}\bar{\phi} \end{pmatrix}.$$

Along a fiber of the cotangent bundle of $\mathcal{S}$, with a fixed local coördinate system, b_{11}, b_{21} remain constant, while ϕ varies. Thus the ratio $b_{11}\phi : b_{21}\phi$ is constant along a fiber of the cotangent bundle ($\phi \neq 0$). This means that every image point of this fiber in S^3, is projected onto the same point of S^2 under the Hopf map. Consequently, the mapping under discussion maps each fiber of $\mathcal{B}$ into a single fiber of the universal bundle. From the nature of the mapping function, it is clear that the map is a homeomorphism of each fiber of $\mathcal{B}$ onto the corresponding fiber of the universal bundle.)

Thus the first Chern class of $\mathcal{B}$ is simply given by $f^*\chi_2$. However, it is well-known [14] that

$$f^*\hat{\psi}_2[\mathcal{S}] = -\text{(Euler characteristic of } \mathcal{S}).$$

Equivalently, we may say that

$$f_*\chi_2 = -(\text{Euler characteristic of } \mathcal{S}) \cdot \hat{\chi}_2,$$

where χ_2, $\hat{\chi}_2$ are the generators of $H_2(\mathcal{S}; Z), H_2(S^2; Z)$, respectively.

We now consider the Gysin homology sequences [15] for the two sphere bundles:

$$\begin{array}{ccccccccc}
0 = H_4(\mathcal{S}) & \to & H_2(\mathcal{S}) & \xrightarrow{j_1} & H_3(\mathcal{B}) & \to & H_3(\mathcal{S}) & = & 0 \\
\downarrow f_* & & \downarrow f_* & & \downarrow \hat{f}_* & & \downarrow f_* & & \\
0 = H_4(S^2) & \to & H_2(S^2) & \xrightarrow{j_2} & H_3(S^3) & \to & H_3(S^2) & = & 0.
\end{array}$$

Because we have a bundle map, the diagram is commutative, and $\hat{f}_* j_1 = j_2 f_*$. The homology sequences are exact, so that j_1, j_2 are isomorphisms onto. Hence $j_1\chi_2 = \chi_3$, $j_2\hat{\chi}_2 = \hat{\chi}_3$, and

$$\begin{aligned}
\hat{f}_* j_1 \chi_2 = \hat{f}_* \hat{\chi}_3 &= j_2 f_* \chi_2 \\
&= -(\text{Euler characteristic}) \cdot j_2\chi_2 \\
&= -(\text{Euler characteristic}) \cdot \chi_3,
\end{aligned}$$

so that the degree of $\hat{f}$, the map obtained from our characteristic matrix is equal to the negative of the Euler characteristic of $\mathcal{S}$.

COROLLARY 7.1. *If $\mathcal{S}$ is an oriented surface of genus $p \neq 1$, and U is a mapping of the cotangent sphere bundle $\mathcal{B}$ of $\mathcal{S}$ into $U(2)$, then*

$$\kappa(U) = -l(U).$$

We now turn our attention to the torus.

THEOREM 7.2. *For an oriented surface $\mathcal{S}$ of genus $p=1$, the characteristic matrix of a special elliptic operator of first order with $P=N=1$, of characteristic ω, has degree 2ω.*

Proof. The characteristic matrix (7.2) may be written in the form

$$(A^1+iA^2)\phi+(A^1-iA^2)\bar{\phi}=\{(A^1T+iA^2T)\phi+(A^1T-iA^2T)\bar{\phi}\}T^{-1}$$

$$=\{A^1T+iA^2T\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\}\begin{pmatrix}\phi & 0\\ 0 & \bar{\phi}\end{pmatrix}T^{-1}+\{A^1T-iA^2T\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\}\begin{pmatrix}\bar{\phi} & 0\\ 0 & \phi\end{pmatrix}T^{-1}.$$

For our special operator,

$$T^{-1}AT=T^{-1}(A^2)^{-1}A^1T=\begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix},$$

so that

$$A^1T-iA^2T\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}=0.$$

Hence the characteristic matrix is

$$\{A^1{}_\alpha T_\alpha+iA^2{}_\alpha T_\alpha\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\}\begin{pmatrix}\phi_\alpha & 0\\ 0_0 & \bar{\phi}_\alpha\end{pmatrix}T_\alpha^{-1}. \tag{7.3}$$

However, the cotangent bundle of $\mathcal{S}$ possesses a non-zero cross section $\{g_\alpha\}$ in the present case, with $g_\alpha dz_\alpha=g_\beta dz_\beta$. The cotangent sphere bundle may then be given by the vectors $g_\alpha e^{i\theta}$ locally. Hence our characteristic matrix defines the function

$$\{A^1{}_\alpha T_\alpha+iA^2{}_\alpha T_\alpha\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\}\begin{pmatrix}g_\alpha & 0\\ 0 & \bar{g}_\alpha\end{pmatrix}\begin{pmatrix}e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{pmatrix}T_\alpha^{-1}$$

on the cotangent sphere bundle $\mathcal{B}$ of $\mathcal{S}$. We now write this function in the form

$$\{A^1{}_\alpha T_\alpha+iA^2{}_\alpha T_\alpha\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\}\begin{pmatrix}g_\alpha & 0\\ 0 & \bar{g}_\alpha\end{pmatrix}T_\alpha^{-1}\cdot T_\alpha\begin{pmatrix}e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{pmatrix}T_\alpha^{-1},$$

i.e., as the product of two matrix functions defined on $\mathcal{B}$. Observe, however, that the first of these comes from a function defined on our non-zero cross section, i.e., a function defined on the torus. This function therefore defines an *inessential* map into S^3 [10, p. 299]. Consequently, the column degree of the first function is zero. The degree of our characteristic matrix is then just

$$l\{T_\alpha\begin{pmatrix}e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{pmatrix}T_\alpha^{-1}\}.$$

In order to compute this number, we must recall the construction of the

special elliptic systems of § 5. C for *general* closed surfaces $\mathcal{S}$. If we use the local coördinate systems which we introduced there, then

$$(7.4)\qquad \begin{pmatrix} \phi_\alpha^{-1} & 0 \\ 0 & \bar{\phi}_\alpha^{-1} \end{pmatrix} T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1}$$

is invariant, since $T \equiv 1$ on $\mathcal{S} - D_\rho$. Since

$$A^1{}_\alpha T_\alpha = iA^2{}_\alpha T_\alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

the characteristic matrix (7.3), may be written as

$$(7.5)\qquad 2A^1{}_\alpha T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0_0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1}.$$

The reasons leading to the invariance of (7.4) also give the invariance of

$$(7.6)\qquad A^1{}_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0_0 & \bar{\phi}_\alpha \end{pmatrix}.$$

From the reasoning employed in Theorem 7.1, we see that the degree of (7.6) is the negative of the Euler characteristic of $\mathcal{S}$. We rewrite (7.4) as

$$\begin{pmatrix} \phi_\alpha^{-1} & 0 \\ 0 & \bar{\phi}_\alpha^{-1} \end{pmatrix} T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1} = \begin{pmatrix} \phi_\alpha^{-1} & 0 \\ 0 & \bar{\phi}_\alpha^{-1} \end{pmatrix} (A^1{}_\alpha)^{-1} \cdot A^1{}_\alpha T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1}.$$

The right side is the product of two function on the cotangent sphere bundle $\mathcal{B}$ of $\mathcal{S}$. If the genus p of $\mathcal{S}$ is *not* equal to 1, then we also know that

$$(7.7)\qquad l\{A^1{}_\alpha T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1}\} = -\kappa\{A^1{}_\alpha T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1}\}$$
$$= 2\omega - 2(1-p),$$

from Theorem 5.2 and Corollary 7.1. On the other hand

$$(7.8)\qquad l\{ \begin{matrix} \phi_\alpha^{-1} & 0 \\ 0 & \bar{\phi}_\alpha^{-1} \end{matrix}\Bigg) (A^1{}_\alpha)^{-1}\} = 2(1-p).$$

Since l is a homomorphism we see that if $p \neq 1$,

$$(7.9)\qquad l\{\begin{pmatrix} \phi_\alpha^{-1} & 0 \\ 0 & \bar{\phi}_\alpha^{-1} \end{pmatrix} T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1}\} = 2\omega.$$

However, when we look at our matrix function (7.4) on the restriction of the cotangent sphere bundle to $\mathcal{S} - D_\rho$, it is simply the constant function 1, so that the bundle is mapped into the single point $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ of S^3 by the first column of the function. Hence the *entire contribution to the degree* must

occur on that part of $\mathcal{B}$ which lies above D_ρ. This, however, is a standard Euclidean neighborhood.

We can now turn our attention to the case $p=1$. Because the cotangent bundle of $\mathcal{S}$ has the non-zero cross section $\{g_\alpha\}$, so that we may take $\phi_\alpha = g_\alpha e^{i\theta}$, the function (7.4) may be written in the form

$$\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} \cdot \left\{ \begin{pmatrix} g_\alpha^{-1} & 0 \\ 0 & \bar{g}_\alpha^{-1} \end{pmatrix} T_\alpha \begin{pmatrix} g_\alpha & 0 \\ 0 & \bar{g}_\alpha \end{pmatrix} T_\alpha^{-1} \right\} \cdot T_\alpha \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} T_\alpha^{-1}.$$

The function

$$\begin{pmatrix} g_\alpha^{-1} & 0 \\ 0 & \bar{g}_\alpha^{-1} \end{pmatrix} T_\alpha \begin{pmatrix} g_\alpha & 0 \\ 0 & \bar{g}_\alpha \end{pmatrix} T_\alpha^{-1} \tag{7.10}$$

comes from a function on the cross section and thus has column degree zero. Clearly,

$$l\left\{ \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} \right\} = 0. \tag{7.11}$$

Hence for the torus,

$$l\left\{ T_\alpha \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} T_\alpha^{-1} \right\} = l\left\{ \begin{pmatrix} \phi_\alpha^{-1} & 0 \\ 0 & \bar{\phi}_\alpha^{-1} \end{pmatrix} T_\alpha \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \bar{\phi}_\alpha \end{pmatrix} T_\alpha^{-1} \right\} \tag{7.12}$$

However, the column degree of the right side depends only on the local behavior of (7.4) above the special neighborhood D_ρ. Hence its degree is 2ω, and

$$l\left\{ T_\alpha \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} T_\alpha^{-1} \right\} = 2\omega.$$

COROLLARY 7.2. *Corollary* 7.1 *remains valid when* $p=1$.

These results immediately yield

THEOREM 7.3. *After a suitable selection of the generator* $\hat{\chi}_3$ *of* $H_3(U(n)/U(1);Z)$

$$\kappa(\mathfrak{a}) = l(\mathfrak{a})$$

for any maping $\mathfrak{a}$ *of the cotangent sphere bundle* $\mathcal{B}$ *of* $\mathcal{S}$ *into* $GL(n,C)$.

UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] M. S. Agranovič, "On the index of elliptic operators," *Doklady Akademii Nauk SSSR*, vol. 142 (1962), pp. 983-985.

[2] M. S. Agranovič and A. S. Dynin, "General boundary value problems for elliptic systems in an n-dimensional domain," *Doklady Akademii Nauk SSSR*, vol. 146 (1962), pp. 1323-1327.

[3] L. Bers, Notes on *Riemann surfaces*, New York University, 1957-8.

[4] A. P. Calderon and A. Zygmund, "Algebras of certain singular operators," *American Journal of Mathematics*, vol. 78 (1956), pp. 310-320.

[5] ———, "Singular integral operators and differential equations," *American Journal of Mathematics*, vol. 79 (1957), pp. 901-921.

[6] S. S. Chern, "An elementary proof of the existence of isothermal parameters on a surface," *Proceedings of the American Mathematical Society*, vol. 6 (1955), pp. 771-782.

[7] N. Dunford and J. T. Schwartz, *Linear Operators*, Part I, Interscience, New York, 1958.

[8] A. S. Dynin, "Singular operators of arbitrary order on a manifold," *Doklady Akademii Nauk SSSR*, vol. 141 (1961), pp. 21-23.

[9] ———, "Multidimensional elliptic boundary value problems for a single unknown function," *Doklady Akademii Nauk SSSR*, vol. 141 (1961), pp. 285-287.

[10] S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, Princeton, 1952.

[11] K. O. Friedrichs, "The identity of weak and strong extensions of differential operators," *Transactions of the American Mathematical Society*, vol. 55 (1944), pp. 132-151.

[12] I. M. Gelfand, "On elliptic equations," *Uspehi Matematičeskih Nauk*, vol. 15 (1960), no. 3, pp. 121-132.

[13] I. C. Gohberg and M. G. Krein, "The basic propositions on defect numbers and indices of linear operators," *American Mathematical Society Translations*, Series 2, vol. 13 (1960), pp. 185-264.

[14] F. Hirzebruch, *Neue topologischen Methoden in der algebraischen Geometrie*, Ergebnisse der Mathematik, Springer, Berlin, 1956.

[15] S. T. Hu, *Homotopy theory*, Academic Press, Nesw York, 1959.

[16] W. Koppelman, "Transmission problems in function theory," *Bulletin of the American Mathematical Society*, vol. 67 (1961), pp. 565-570.

[17] M. Schechter, "Remarks on elliptic boundary value problems," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 561-578.

[18] M. Schiffer and D. C. Spencer, *Functionals of Finite Riemann Surfaces*, Princeton, 1954.

[19] L. Schwartz, *Ecuaciones Differenciales Parciales Elipticas*, Bógota, 1956.

[20] R. T. Seeley, "Singular integrals on compact manifolds," *American Journal of Mathematics*, vol. 81 (1959), pp. 658-690.

[21] R. T. Seeley, "Regularizations of singular integral operators on compact manifolds," *American Journal of Mathematics*, vol. 83 (1961), pp. 265-275.

[22] N. Steenrod, *The Topology of Fibre Bundles*, Princeton, 1951.

[23] A. I. Volpert, "The index of boundary value problems for systems of harmonic functions in three independent variables," *Doklady Akademii Nauk SSSR*, vol. 133 (1960), pp. 13-15.

[24] A. I. Volpert, "On the index of systems of two-dimensional singular integral equations," *Doklady Akademii Nauk SSSR*, vol. 142 (1962), pp. 776-777.

INTEGRAL EQUIVALENCE OF VECTOR FIELDS ON MANIFOLDS AND BIFURCATION OF DIFFERENTIAL SYSTEMS.*

By A. Aeppli and L. Markus[1].

1. Basic equivalence relations of vector fields and examples. Two vector fields on a differentiable manifold M^n are integrally equivalent if their integral curve families are topologically equivalent, see below for precise definitions. We shall study the relation of integral equivalence of vector fields on M^n especially with reference to the concepts of structural stability and bifurcation of differential systems. We shall use the techniques of homotopy, for nowhere vanishing vector fields, as a guide in the theory but we note that a single homotopy class can contain a noncountable number of integral equivalence classes. We apply our general theory to a detailed analysis of nowhere vanishing vector fields on the torus surface T^2 and we find important properties and canonical representatives of the various integral equivalence classes. The existence of a noncountable number of integral equivalence classes for nowhere vanishing vector fields on the torus T^n is used to demonstrate the existence of noncountably many such integral classes on higher dimensional manifolds.

Let M^n be a C^∞ differentiable n-manifold, that is, M^n is a connected separable metrizable space which is locally homeomorphic to the real number space R^n, $n \geqq 1$, and the local coordinates so determined are inter-related by C^∞ maps with non vanishing jacobians. A first order, autonomous, real, ordinary differential system $\mathcal{S}$ on M^n is a tangent vector field v on M^n, that is, a cross-section of M^n into its contravariant tangent bundle $T(M^n)$. In local coordinates $(x^1, x^2, \cdots, x^n)$ on M^n we write the differential system $\mathcal{S}$ as

$$\mathcal{S}: \dot{x}^i = v^i(x^1, \cdots, x^n) \qquad i = 1, \cdots, n.$$

Let $\mathfrak{V}(M^n)$ be the real linear space of all C^1 tangent vector fields on M^n with the usual compact-open C^1-topology. We define this topology on $\mathfrak{V}(M^n)$ by a metric[2]

* Received February 28, 1963.

[1] Research supported by NONR 3776(00).

[2] We put $Dv = \operatorname{grad} v = \left(\frac{\partial v^i}{\partial x^j}\right)$, and the norms are defined by $|v| = \sup_i |v^i|$, $|Dv| = \sup_{i,j} \left|\frac{\partial v^i}{\partial x^j}\right|$.

$$\rho(v_1, v_2) = \sum_{\alpha=1}^{\infty} 1/2^{\alpha} \sup \frac{|v_1 - v_2| + |Dv_1 - Dv_2|}{1 + |v_1 - v_2| + |Dv_1 - Dv_2|}$$

where each term of the sum is the supremum in a local coordinate patch U_α, $\alpha = 1, 2, \cdots$, the sets U_α cover M^n, and each $\bar{U}_\alpha$ is compact and lies within some local coordinate patch on M^n. The topology so defined on $\mathfrak{V}(M^n)$ is independent of the choice of the covering U_α. Thus $\mathfrak{V}(M^n)$ is a Fréchet space and if M^n is compact it is a Banach space. We shall often deal with the subset $\mathfrak{V}_0(M^n)$ consisting of all vector fields in $\mathfrak{V}(M^n)$ which vanish nowhere on M^n. If M^n is compact, $\mathfrak{V}_0(M^n)$ is an open subset of $\mathfrak{V}(M^n)$.[3]

Definition. Vector fields v_0 and v_1 in $\mathfrak{V}(M^n)$ are integrally equivalent in case there exists a homeomorphism of M^n onto itself which carries the sensed (but not parametrized) integral curve family of v_0 onto that of v_1.

Here an integral curve $\phi(t)$ of $v \in \mathfrak{V}(M^n)$ is a maximal solution curve of the corresponding differential system

$$\mathcal{S}: \dot{x}^i = v^i(x^1, \cdots, x^n) \qquad i = 1, \cdots, n$$

and a sensed integral curve is the class of all curves $\phi(t(\tau))$ where $\tau \to t(\tau)$ is an orientation preserving homeomorphism of R' onto itself. Thus if v_0 and v_1 in $\mathfrak{V}(M^n)$ are integrally equivalent by a homeomorphism Φ, then Φ maps each sensed integral curve of v_0 onto one of v_1 and vice versa for Φ^{-1}, and Φ maps critical points, periodic solutions, invariant and in particular minimal sets of v_0 onto the corresponding objects of v_1.

If v_0 and v_1 lie in $\mathfrak{V}_0(M^n)$ we say they are homotopic in case the corresponding cross-sections in $T_0(M^n)$, the bundle of nowhere vanishing tangent vectors, are homotopic maps by a continuous homotopy

$$V: M^n \times I \to T_0(M^n),$$

which is a cross-section for each fixed t in the interval $I: 0 \leq t \leq 1$.

PROPOSITION. *Vector fields v_0 and v_1 in $\mathfrak{V}_0(M^n)$ are homotopic if and only if they lie in the same arc-component of $\mathfrak{V}_0(M^n)$. If M^n is compact, each component of $\mathfrak{V}_0(M^n)$ is also an arc-component.*

Proof. The vector fields v_0 and v_1 are in the same arc-component of

[3] C^1 continuity for vector fields has to be considered if stability problems are discussed whereas a strong form of C^0 continuity (Lipschitz) is usually enough in dealing with questions on integral equivalence (i.e. C^0 integral equivalences as defined here; there are pairs of integrally equivalent vector fields one of which is structurally stable, the other one unstable, cf. [7] and see Example 3). Thus Theorems 1, 2, 3, 4, 5, 8 make sense if the C^0 topology (and Lipschitz continuity) is used.

$\mathcal{V}_0(M^n)$ just in case there is a continuous map of the interval $I: 0 \leqq t \leqq 1$ into $\mathcal{V}_0(M^n)$ which defines a curve v_t joining v_0 to v_1. In this case

$$(P, t) \to v_t(P)$$

defines a map

$$M^n \times I \to T_0(M^n)$$

which is a C^1 cross-section for each fixed $t \in I$ and which is continuous on the product $M^n \times I$. Hence v_0 and v_1 are homotopic.

On the other hand let v_0 and v_1 in $\mathcal{V}_0(M^n)$ be homotopic by

$$V: M^n \times I \to T_0(M^n).$$

We define a cross-section

$$V \times 1: M^n \times I \to T_0(M^n) \times I$$

by

$$V \times 1: (P, t) \to (V(P, t), t)$$

for $P \in M^n$ and $t \in I$. We approximate the continuous cross-section $V \times 1$ by a C^1-cross-section [12] and so construct a C^1-homotopy V^1 of v_0 to v_1

$$V^1: M^n \times I \to T_0(M^n): (P, t) \to \hat{v}_t(P).$$

Let $U: (x^1, \cdots, x^n)$ be a local coordinate system on M^n with compact $\bar{U}$ lying in some local coordinate system, and we write

$$V^1: \dot{x}^i = \hat{v}_t^i(x^1, \cdots, x^n) = \hat{v}^i(x, t), \qquad i = 1, \cdots, n,$$

where $\hat{v}^i(x^1, \cdots, x^n, t)$ is in C^1 in $U \times I$. Thus

$$\sup_{x \in U;\, i, j} [| \hat{v}^i(x, t_1) - \hat{v}^i(x, t_2) | + | \frac{\partial \hat{v}^i}{\partial x^j}(x, t_1) - \frac{\partial \hat{v}^i}{\partial x^j}(x, t_2) |] < \epsilon$$

whenever $| t_2 - t_1 | < \delta(\epsilon)$, for a prescribed $\epsilon > 0$.

Choose a sequence $U_1, U_2, \cdots, U_\alpha, \cdots$ of such coordinate systems covering M^n and we then compute the distance in $\mathcal{V}_0(M^n)$,

$$\rho(v_{t_2}, v_{t_1}) \leqq \sum_{\alpha=1}^{N} \frac{1}{2^\alpha} \frac{\epsilon}{1+\epsilon} + \sum_{\alpha=N+1}^{\infty} 1/2^\alpha$$

whenever $| t_2 - t_1 | < \min [\delta_1(\epsilon), \cdots, \delta_N(\epsilon)]$. Thus the map

$$I \to V_0(M^n): t \to \hat{v}_t$$

is continuous and so v_0 and v_1 are in the same arcwise connected component of $\mathcal{V}_0(M^n)$.

Now let M^n be compact. Then $\mathcal{V}(M^n)$ is a Banach space and $\mathcal{V}_0(M^n)$ is an open subset of $\mathcal{V}(M^n)$. Thus $\mathcal{V}_0(M^n)$ is locally arcwise connected and

each component of $\mathcal{V}_0(M^n)$ is open in $\mathcal{V}(M^n)$. Let A be a component of $\mathcal{V}_0(M^n)$ and let $v \in A$. The set of points of $\mathcal{V}_0(M^n)$ which can be joined to v by a continuous curve in $\mathcal{V}_0(M^n)$ is clearly open and closed in A and hence must coincide with A. Thus A is arcwise connected. Q. E. D.

In general, integrally equivalent vector fields v_0 and v_1 of $\mathcal{V}_0(M^n)$ are not homotopic. However if the integral equivalence is defined by a homeomorphism Φ which is sufficiently "close to the identity and smooth," then v_0 and v_1 are homotopic, as is shown next.

PROPOSITION. *Let v_0 and v_1 in $\mathcal{V}_0(M^n)$ be integrally equivalent by the homeomorphism Φ of M^n onto itself. If Φ is C^1-isotopic to the identity map of M^n, then v_0 and v_1 are homotopic.*

Proof. Let

$$\Phi_t \colon M^n \times I \to M^n \colon (P, t) \to \Phi_t(P)$$

be the isotopy with $\Phi_t(P)$ in C^1 on $M^n \times I$ and

$$\Phi_1(P) = \Phi(P), \qquad \Phi_0(P) = P.$$

Here $\Phi = \Phi_1$ carries the integral curve family of v_0 onto that of v_1. Since each Φ_t is a C^1-diffeomorphism of M^n onto itself, we can define the vector field $v_t \in \mathcal{V}_0(M^n)$ as the image of v_0 under the map Φ_t. Then v_t is the required homotopy joining v_0 to v_1. Q. E. D.

These ideas of homotopy and of integral equivalence have applications to the concepts of structural stability and bifurcation in the theory of differential equations, as described below.

A differential system, or vector field $v \in V(M^n)$, is structurally stable in case: for each prescribed neighborhood H of the identity in the group of all homeomorphisms of M^n onto itself, there exists a neighborhood N of v in $\mathcal{V}(M^n)$ such that each $v_1 \in N$ is integrally equivalent to v by a homeomorphism of H, see [7]. Thus for a structurally stable $v \in \mathcal{V}(M^n)$, the integral equivalence class of v contains a full neighborhood of v in $\mathcal{V}(M^n)$.

A continuous map from a topological space Λ into $\mathcal{V}(M^n)$

$$\Lambda \to \mathcal{V}(M^n) : \lambda \to v_\lambda \in \mathcal{V}(M^n)$$

is called a family of differential systems which is parametrized by Λ. In local coordinates $(x^1, \cdots, x^n)$ on M^n we write

$$\mathcal{S}_\lambda \colon \dot{x}^i = v_\lambda{}^i(x^1, \cdots, x^n) = v^i(x, \lambda) \qquad i = 1, \cdots, n$$

where $v^i(x, \lambda)$ and $\partial v^i / \partial x^j(x, \lambda)$ are continuous in $(x) \times \Lambda$. A point $\lambda_0 \in \Lambda$ is an ordinary point for the parametrized family $\mathcal{S}_\lambda$ in case there exists a neighborhood L of λ_0 in Λ such that $\mathcal{S}_{\lambda_1}$ is integrally equivalent to $\mathcal{S}_{\lambda_0}$ for each $\lambda_1 \in L$. If λ_0 is not an ordinary point, then λ_0 is called a bifurcation

point for the parametrized family $\mathcal{S}_\lambda$. Clearly the set of ordinary points of $\mathcal{S}_\lambda$ is open and the set of bifurcation points is closed in Λ.

Example 1. Consider the n-torus T^n as a differentiable manifold obtained from R^n modulo the group of integral translations. Using coordinates $(x^1,\cdots,x^n)$ (mod 1), consider the differential systems on T^n

$$\mathcal{S}_\lambda : \dot{x}^i = \lambda^i \qquad\qquad i=1,\cdots,n$$

where $\lambda=(\lambda^1,\cdots,\lambda^n)$ is a constant nonzero n-vector. Then $\mathcal{S}_\lambda$ is a family of differential systems in $\mathfrak{V}_0(T^n)$, parametrized by $\Lambda = R^n - 0$. If the vector space F_λ generated by $\lambda^1,\cdots,\lambda^n$ over the rational field F has dimension k, $1\leqq k\leqq n$, then each minimal set (compact invariant set which contains no proper compact invariant subset) of $\mathcal{S}_\lambda$ is a k-torus, see [3]. In particular every solution curve of $\mathcal{S}_\lambda$ is periodic if and only if $k=1$. Since the rational vectors λ are dense in Λ, none of the differential systems $\mathcal{S}_\lambda$ is structurally stable and every value of $\lambda\in\Lambda$ is a bifurcation value.

Choose a homology basis for T^n, say the cycles corresponding to the coordinate circles of $x^1,x^2,\cdots,x^n$. Let $S(t)$ be an integral curve of $\mathcal{S}_\lambda$ initiating at $P\in T^n$ and let N be a small ball neighborhood of P in T^n. Let $t_1<t_2<\cdots<t_r<\cdots$ be an infinite sequence of times for which $S(t_r)$ meets N, $\lim_{r\to\infty} t_r=\infty$, and let $\ell_1,\ell_2,\cdots,\ell_r,\cdots$ be the sequence of homology classes in $H_1(T^n)$ determined by $S(t)$, $0\leqq t\leqq t_r$, and thence joining $S(t_r)$ with P by a small arc in N. If we express ℓ_r in terms of the chosen basis of $H_1(T^n)$ we obtain a nonzero integral vector

$$\ell_r = (a_r^1,\cdots,a_r^n).$$

As $r\to\infty$ the direction cosines of the vector l_r approach $\lambda^i/|\lambda|$, $i=1,\cdots,n$, where $|\lambda| = (\lambda^{1^2}+\cdots+\lambda^{n^2})^{\frac{1}{2}}$. Thus the homogeneous coordinates $(\lambda^1,\cdots,\lambda^n)$ are topological invariants of $\mathcal{S}_\lambda$ and are defined to be the generalized rotation number of $\mathcal{S}_\lambda$ relative to the given basis of $H_1(T^n)$, cf. [11].

A homeomorphism of T^n onto itself induces an automorphism of $H_1(T^n)$ which is described by a unimodular matrix A (integer matrix with determinant ± 1), relative to a basis of $H_1(T^n)$. Thus the generalized rotation number of $\mathcal{S}_\lambda$ with respect to any basis of $H_1(T^n)$ is a unimodular transform of λ, that is, $A\lambda$. Since there is only a countable number of unimodular matrices, there is at most a countable number of the systems $\mathcal{S}_\lambda$ which are integrally equivalent to a fixed $\mathcal{S}_{\lambda_0}$. Therefore there is a continuum of integral equivalence classes of $\mathfrak{V}_0(T^n)$ represented by the various systems $\mathcal{S}_\lambda$ for $\lambda\in\Lambda$.

In the study of $\mathfrak{V}_0(T^2)$ we can define the classical (non-homogeneous) rotation number λ of a nonlinear differential system [2]

$$\mathcal{S}: \dot{x}^1 = 1, \qquad \dot{x}^2 = g(x^1, x^2)$$

using a point P on a recurrent solution $S(t)$ of $\mathcal{S}$ on T^2. The rotation number λ, defined just as above, is independent of the point P, the neighborhood N, and the recurrent solution $S(t)$. Hence λ depends only on $\mathcal{S}$ and the homology basis of $H_1(T^2)$. There is a periodic solution of $\mathcal{S}$ if and only if λ is rational.

THEOREM 1. *Let M^n be a differentiable manifold. The cardinal number of integral equivalence classes in $\mathfrak{V}(M^n)$ is that of the continuum. If $n \geqq 2$ and $\mathfrak{V}_0(M^n)$ is not empty, then there is a continuum of integral equivalence classes in $\mathfrak{V}_0(M^n)$.*

Proof. Since $\mathfrak{V}(M^n)$ has the cardinality $\mathfrak{c}$ of the continuum, the number of integral equivalence classes is at most $\mathfrak{c}$.

Let $v_0 \in \mathfrak{V}(M^n)$ be a vector field with precisely one zero P_0 on M^n. Take local coordinates on M^n which define a diffeomorphism of an open set $U \subset M^n$ with the ball $x^{1^2} + \cdots + x^{n^2} < 1$ in R^n, with P_0 corresponding to the origin. Let f be a real differentiable function on M^n such that $f(P) = 1$ on $M^n - U$ and $f(x) = 1$ for $x^{1^2} + \cdots + x^{n^2} \geqq \frac{1}{4}$. Let $Z \subset U$ be the compact set of zeros of f and assume $P_0 \in Z$. Then Z is the set of critical points of fv_0 on M^n. But the cardinality of the number of homeomorphism types of compact subsets of a ball is certainly $\mathfrak{c}$. Thus the vector fields $fv_0 \in \mathfrak{V}(M^n)$, as f describes the admissible class of real functions, determines a continuum of distinct integral equivalence classes.

Now let $n \geqq 2$ and assume that M^n is either noncompact or M^n is compact with Euler characteristic zero so that $\mathfrak{V}_0(M^n)$ is not empty. First assume $n \geqq 3$ and let T^{n-1} be a small torus submanifold contained entirely in one local coordinate system $(x^1, \cdots, x^n)$ of M^n. We define a vector field $v_0 \in \mathfrak{V}(M^n)$ so that T^{n-1} is a minimal set for v_0, and on T^{n-1} the vector field is integrally equivalent to a linear differential system $\mathcal{S}_\lambda$, as in example 1. This is possible since such a field exists on the solid torus A bounded by T^{n-1}, and since M^n is either noncompact or has Euler characteristic zero. By appropriate choice of λ we can describe a continuum of topologically inequivalent vector fields on T^{n-1}. In order to complete the proof we must define $v_0 \in \mathfrak{V}_0(M^n)$ so that any homeomorphism of M^n onto itself which preserves the integral curve family of v_0, must map T^{n-1} onto itself.

For convenience use a real analytic structure on M^n with a real analytic Riemann metric (this can be done by Whitney's imbedding theorem). We modify v_0 near T^{n-1} so that T^{n-1} is stable, that is, the normal component of v_0 points towards T^{n-1} from both sides of T^{n-1}. Now approximate v_0 by a real analytic vector field v_1 outside a neighborhood of the solid torus A which is

bounded by T^{n-1}. Using interpolation to blend v_1 into v_0, we obtain a vector field $v_2 \in \mathfrak{V}_0(M^n)$ where $v_2 = v_1$ outside a neighborhood of A, v_2 has normal component directed towards T^{n-1} near T^{n-1}, and $v_2 = v_0$ on T^{n-1}. Inside T^{n-1} we can further require that v_2 has a family of periodic solutions, with the same period, which fill a subset Σ of A so that Σ is compact and has a non-empty interior in A. Since v_2 is real analytic outside a neighborhood of A, it has no collection of periodic solutions homeomorphic to Σ outside A. Therefore T^{n-1} can be topologically characterized as the unique minimal $(n-1)$-torus which contains within it a family of periodic solutions whose union is compact and also has a non empty interior in M^n.

Finally let $n = 2$. If M^2 is compact with Euler characteristic zero, then M^2 is either the torus T^2 or the Klein bottle. In either case it is easy to construct a continuum of topologically different vector fields in $\mathfrak{V}_0(M^2)$ by demanding topologically different sets of periodic solutions on M^2.

If M^2 is a noncompact surface we consider the boundary B of M^2 (defined by sequences of compact subsets of M^2 under exhaustion, see [1, 4]). If B has just one component, then M^2 is either the punctured sphere, the projective plane, or the Klein bottle; or M^2 has a handle. If M^2 is R^2, then there is a continuum of integral equivalence classes in $\mathfrak{V}_0(R^2)$, as is shown in [6]. If M^2 is the punctured projective plane or the Klein bottle, it is easy to give explicit descriptions for a continuum of integral equivalence classes. If M^2 has a handle, then we can construct two smooth simple closed curves around the meridians of the handle, to form an annular region A around the handle. Then define a continuum of topologically different structures of periodic solutions within A and define the corresponding vector field outside A to be real analytic and nonvanishing. In this way we can define a continuum of integral equivalence classes of $\mathfrak{V}_0(M^2)$.

If the boundary B of the noncompact surface M^2 is not connected then let A be an annulus bounded by two smooth closed curves on M^2 which separate two components of B. Then a vector field v defined on A with a prescribed pattern of periodic solutions, can be extended to all M^2 with no critical points. Here again we obtain a continuum of integral equivalence classes in $\mathfrak{V}_0(M^2)$. Q. E. D.

Example 2. Consider T^2 as R^2 modulo the group of integral translations so that (x^1, x^2) (mod 1) serve as coordinates on T^2. Consider the differential systems v_1 and $v_{-1} \in \mathfrak{V}_0(T^2)$,

$$v_1: \dot{x}^1 = \cos 2\pi x^2, \qquad \dot{x}^2 = \sin 2\pi x^2$$

and

$$v_{-1}: \dot{x}^1 = -\cos 2\pi x^2, \qquad \dot{x}^2 = \sin 2\pi x^2.$$

The homeomorphism

$$\Phi : (x^1, x^2) \to (-x^1, x^2)$$

defines an integral equivalence of v_1 with v_{-1}. But the degree of v_1 is $(1, 0)$ and the degree of v_{-1} is $(-1, 0)$, computed with respect to the cycles $x^1 = 0$ and $x^2 = 0$ as defined below. Hence v_1 and v_{-1} are not homotopic.

Example 3. Consider in R^2

$$v : \dot{x}^1 = \cos x^2, \qquad \dot{x}^2 = \sin x^2$$

which has the lines $x^2 = 0$, $x^2 = n\pi$ for $n = \pm 1, \pm 2, \cdots$ as integral curves. Using the integral curves of v, we can define a map $f : y_0 \to y(y_0)$ from initial points $(0, y_0)$ for $y_0 \to 0+$, to endpoints $(0, y)$ for $y \to \pi-$. The map f is nonsingular in that $\lim_{y \to 0+} dy/dy_0 \neq 0$, in fact this derivative is nonzero regardless of the differentiable local coordinates used near $(0, 0)$ and $(0, \pi)$ in R^2. Now we define a homeomorphism of R^2 onto itself by

$$\Phi : x^1 \to \bar{x}^1 = x^1, \qquad x^2 \to \bar{x}^2 = (x^2)^3.$$

Then in R^2 we define the system

$$\bar{v} : \dot{x}^1 = \cos (x^2)^{1/3}, \qquad \dot{x}^2 = 3 (x^2)^{2/3} \sin (x^2)^{1/3}.$$

We note that v and $\bar{v}$ are integrally equivalent by the homeomorphism Φ. However v and $\bar{v}$ are not integrally equivalent by a C^1-diffeomorphism of R^2 onto itself since for $\bar{v}$ the map $\bar{f}$, defined just as the map f for v, must be singular.

This example shows that the relation of integral equivalence under C^1-homeomorphisms is too delicate for the geometric analysis of the qualitative nature of differential systems. It is easy to construct examples of differential systems v and $\bar{v}$ in $\mathfrak{V}_0(T^2)$ which are topologically, but not C^1-homeomorphically, integrally equivalent. We need only construct a v with two periodic solutions around meridians of T^2, each with a non-vanishing characteristic exponent. Then take $\bar{v}$ to be topologically integrally equivalent to v but require that all characteristic exponents of $\bar{v}$ are zero.[4]

2. Integral equivalence on the torus T^2. Consider the torus T^2 as a differentiable (this means C^∞ unless otherwise qualified) manifold. Choose a diffeomorphism of T^2 with $S^1 \times S^1$, where S^1 is the circle group of all complex numbers of unit modulus. Use the parallelization defined on T^2 by this product structure to fix a diffeomorphism of the tangent bundle $T(T^2)$

[4] Similar examples exist in higher dimensions.—In dimension 2, the characteristic exponent is invariant under C^1 diffeomorphisms; in general, the transversal germ of a periodic solution is a linear transformation $R^{n-1} \to R^{n-1}$ whose eigenvalues (the characteristic multipliers) are C^1 diffeomorphism invariants. Cf. [7].

with $T^2 \times S^1$. For each vector field $v \in \mathcal{V}_0(T^2)$ we then have a well-defined map $T^2 \to T^2 \times S^1$. If we write T^2 as $S^1 \times S^1$ and use the projection map of $T^2 \times S^1$ onto S^1 we obtain a map

$$S^1 \times S^1 \to S^1.$$

This map, restricted to $1 \times S^1$ or to $S^1 \times 1$, has an integral degree.

Definition. Fix a diffeomorphism $T^2 = S^1 \times S^1$, and then for each vector field $v \in V_0(T^2)$ the map $S^1 \times S^1 \to S^1$ is defined with degrees a_1 and a_2 for the restriction to each factor, as above. The degree of v is the ordered pair of integers (a_1, a_2) with respect to the assumed product structure of T^2.

Remark. The degree (a_1, a_2) of $v \in \mathcal{V}_0(T^2)$ depends on the diffeomorphism $T^2 = S^1 \times S^1$. in particular on the basis of cycles for $H_1(T^2)$, the two basic cycles being parametrized by the angles θ_1, θ_2 such that we get a parametrization of $S^1 \times S^1$. If this diffeomorphism is replaced by another, obtained by composition with an automorphism of $S^1 \times S^1$,

$$\theta_1 \to \bar{\theta}_1 = \alpha\theta_1 + \beta\theta_2$$
$$\theta_2 \to \bar{\theta}_2 = \gamma\theta_1 + \delta\theta_2$$

where the integral matrix

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

has determinant

$$\det A = \pm 1,$$

then the degree of v becomes $(\bar{a}_1, \bar{a}_2)$ where

$$\bar{a}_1 = a_1\alpha + a_2\beta$$
$$\bar{a}_2 = a_1\gamma + a_2\delta.$$

Since any diffeomorphism of $S^1 \times S^1$ onto itself is differentiably homotopic to an automorphism of $S^1 \times S^1$ as above [4, 12], the degree of v can be (a_1, a_2) or any pair of integers

$$\begin{pmatrix} \bar{a}_1 \\ \bar{a}_2 \end{pmatrix} = A \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

where the integral matrix A has determinant $+1$ or -1, and all such unimodular matrices A can occur.

THEOREM 2. *Fix a diffeomorphism $T^2 = S^1 \times S^1$ and consider the collection $\mathcal{V}_0(a, b)$ of vector fields in $\mathcal{V}_0(T^2)$ with degree (a, b). Let*

$$d = \text{g.c.d.}(a, b) > 0.^5$$

Then there exists a diffeomorphism of T^2 onto itself which carries all members of $\mathcal{V}_0(a, b)$ onto all members of $\mathcal{V}_0(d, 0)$. Furthermore if

$d_1 \neq d_2$ are non negative integers

then any $v_1 \in \mathfrak{V}_0(d_1, 0)$ and $v_2 \in \mathfrak{V}_0(d_2, 0)$ are not integrally equivalent.

Proof. To prove the existence of a diffeomorphism of T^2 onto itself carrying $\mathfrak{V}_0(a, b)$ onto $\mathfrak{V}_0(d, 0)$ it is sufficient to construct a unimodular matrix A such that

$$\begin{pmatrix} d \\ 0 \end{pmatrix} = A \begin{pmatrix} a \\ b \end{pmatrix}$$

or

$$a\alpha + b\beta = d$$
$$a\gamma + b\delta = 0.$$

Now a/d and b/d are relatively prime integers and so there exist integers α and β with

$$\alpha(a/d) + \beta(b/d) = 1.$$

Take

$$\gamma = -b/d, \qquad \delta = a/d$$

so

$$\alpha\delta - \beta\gamma = \alpha(a/d) + \beta(b/d) = 1.$$

To prove that no members of $\mathfrak{V}_0(d_1, 0)$ and $\mathfrak{V}_0(d_2, 0)$ are integrally equivalent, for $d_1 \neq d_2$, note that any such homeomorphism of T^2 carrying the integral curves of $v_1 \in \mathfrak{V}_0(d_1, 0)$ onto those of $v_2 \in \mathfrak{V}_0(d_2, 0)$ can be factored into a diffeomorphism Φ followed by a homeomorphism Ψ homotopic to the identity. Now Φv_1 has degree, relative to the fixed product structure,

$$A \begin{pmatrix} d_1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix},$$

where A is an integral matrix with determinant ± 1. Finally we shall prove that the degree of $\Psi\Phi v_1 = v_2$ is also (a, b).

Let ℓ_1 and ℓ_2 be homeomorphic images of S^1 in T^2 which are homotopic to the curves $1 \times S^1$ and $S^1 \times 1$ in $T^2 = S^1 \times S^1$, respectively. If ℓ_1 is a nonsingular differentiable curve with parameter θ_1 on $0 \leqq \theta_1 < 2\pi$, let $\Delta\theta_1$ be the angle between the tangent vector to ℓ_1 and the vector of v_2, in the positive sense of the orientation of T^2. Then

$$d_2 = \tfrac{1}{2\pi} \int d\Delta\theta_1$$

where the integral is computed positively around ℓ_1. A similar integral around ℓ_2 would yield zero. Now we can define this integral around ℓ_1 purely topologically. Assume ℓ_1 is only a topological image of S^1 which is homotopic to

[5] g. c. d. $(0, 0) = 0$ by definition.—$\mathfrak{V}_0(a, b)$ are the connectivity components of $\mathfrak{V}_0(T^2)$, $-\infty < a, b < +\infty$, corresponding to the elements of the Bruschlinsky group $\pi(S^1 \times S^1, S^1) \cong H^1(S^1 \times S^1)$.

$1 \times S^1$. Assume further that the integral curve family of v_2 has only a finite number of simple crossings with ℓ_1. That is, at every point of ℓ_1 except at the crossing points the corresponding integral curve of v_2 initially penetrates one of the two sides of ℓ_1 in T^2, and at the finite number of crossing points of ℓ_1 the designated side changes. Also assume that, at a point which is not crossing point, a short segment of an integral curve of v_2 crosses a sufficiently small circle (around the intersection point of ℓ_1 with the considered integral curve of v_2) either ahead of the corresponding point of ℓ_1 or else definitely behind the corresponding point of ℓ_1. In this way we can define positively rotating and negatively rotating crossings of v_2 over ℓ_1. If these crossings are counted $(+1)$ and (-1) respectively, the algebraic sum of the crossings of v_2 over ℓ_1 is exactly d_2, see [4]. A similar construction for ℓ_2 yields the crossing sum of 0.

Now under the homeomorphism Ψ^{-1} of T^2 carrying v_2 onto Φv_1, the crossing sums are invariant, since Ψ^{-1} is homotopic to the identity. That is, the crossing sum of v_2 over ℓ_1 is equal to that of Φv_1 over $\Psi^{-1}\ell_1$, which shows that

$$d_2 = a.$$

Similarly

$$0 = b.$$

But there is no unimodular matrix A which transforms the vector $\begin{pmatrix} d_1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} d_2 \\ 0 \end{pmatrix}$ if $d_1 \neq d_2$. Therefore $v_1 \in \mathfrak{V}_0(d_1, 0)$ is not integrally equivalent to $v_2 \in \mathfrak{V}_0(d_2, 0)$ when $d_1 \neq d_2$. Q. E. D.

COROLLARY. *If $v \in \mathfrak{V}_0(T^2)$ has degree $(0,0)$ in one product structure of T^2 then v also has the same degree $(0,0)$ with respect to any other product structure $T^2 = S^1 \times S^1$ of the torus.*

Remark. From the result of Theorem 2 we can obtain a complete list of all integral equivalence classes in $\mathfrak{V}_0(T^2)$ by listing the integral equivalence classes in each of the separate spaces $\mathfrak{V}_0(0,0), \mathfrak{V}_0(1,0), \mathfrak{V}_0(2,0), \cdots$.

The next theorem cites a geometric property of the set $\mathfrak{V}_0(d,0)$ which emphasizes the significance of the non-negative integer d. We will count the number of periodic solutions disregarding differences in periodic solutions due only to translations in time.

THEOREM 3. *If $v \in \mathfrak{V}_0(d,0)$ for $d \geqq 0$, then there are at least $2d$ periodic solutions of v on T^2. There exist members of $\mathfrak{V}_0(d,0)$ which have exactly $2d$ periodic solutions.*

Proof. Suppose there are no periodic solutions in v. Then using the

orthogonal trajectory to v, with a slight modification to construct a closed curve, see [2], we can obtain a C^1 diffeomorphism of T^2 onto itself so that v has positive longitudinal velocity in the new product coordinates. Call these coordinates (x,y) (mod 1) and around the meridian circle $x=0$ we note the degree $a=0$, since $\dot{x}>0$ on v. To compute the degree b around the longitude circle $y=0$ we use instead of this circle a homotopic curve composed of an arc of a solution of v between two consecutive intersections with $x=0$ and also a segment of the meridian $x=0$. We note that the degree of v on this closed curve, as computed by the line integral in Theorem 2, is zero. Thus v has degree $(0,0)$ in the new product coordinates on T^2 and hence also in the original product coordinates. Thus $v\in\mathfrak{V}_0(0,0)$.

Let $v\in\mathfrak{V}_0(d,0)$ with $d\geqq 1$. Then v has at least one periodic solution; and we assume that v has only a finite number of periodic solutions. Suppose there were just one periodic solution S of v. Then there is a C^1-homeomorphism of T^2-S onto a plane annulus A and we can identify both boundary curves of A with S. Each solution curve of v', the image of v in A, must have its negative limit set on one boundary curve of A and its positive limit set on the other boundary curve. For otherwise the Poincaré-Bendixson theorem guarantees the existence of a periodic solution of v' interior to A and hence a periodic solution of v on T^2 different from S. Now we can compute the degree of v on T^2 relative to a homology basis of closed curves ℓ_1, ℓ_2. Take ℓ_1 to be the periodic solution S which is homotopic to the longitude circle on T^2. For ℓ_2 use an arc of the spiral solution of v which approaches S from one side as $t\to-\infty$ and S from the other side as $t\to+\infty$ and then supplement this arc by two short pieces of the orthogonal trajectory of v to reach S so as to make ℓ_2 a closed curve. Using the line integral described in Theorem 2, we compute the degree of v around ℓ_2 to be zero. Thus the degree of v relative to some product structure is $(0,0)$, which contradicts the assumption that $v\in\mathfrak{V}_0(d,0)$ with $d\geqq 1$. Thus v must have at least two periodic solutions.

Therefore the first assertion of the theorem is proved for $d=0$ (trivially) and $d=1$. Now assume $d\geqq 2$ and $v\in\mathfrak{V}_0(d,0)$ has a finite number of periodic solutions. There exists a C^∞ nonsingular simple closed curve $\mathcal{B}$, homotopic to the meridian circle on T^2, such that $\mathcal{B}$ meets each periodic solution of v in just one point and each such intersection is orthogonal in the obvious Riemannian metric on $T^2=S^1\times S^1$. Let $\Delta\theta$, as a function of the angular coordinate θ on $\mathcal{B}$, describe the angle from the tangent vector to $\mathcal{B}$ positively around to the vector field v, as in Theorem 2.

Let S_1 and S_2 be consecutive periodic solutions of v as ordered by their intersection points with $\mathcal{B}$. Then

$$\frac{1}{2\pi}\int d\Delta\theta(\theta) \qquad \text{from } S_1\cap\mathcal{B} \text{ to } S_2\cap\mathcal{B}$$

is $p/2$ for some integer p. By considering the region of T^2 bounded by S_1 and S_2 as a plane annulus and using the Poincaré-Bendixson theorem as above, we can prove that $p=0$ or $p=\pm 1$. Since each interval along $\mathcal{B}$ between consecutive periodic solutions of v can contribute only 0, $\frac{1}{2}$, or $-\frac{1}{2}$ to the line integral for the degree of v around $\mathcal{B}$, there must be at least $2d$ such intervals. Thus there must be at least $2d$ periodic solutions of v on T^2.

Finally we give examples of vector fields which realize the minimum number of periodic solutions for each $d\geqq 0$.

For $d=0$ consider

$$\dot{x}=1, \qquad \dot{y}=\sqrt{2}$$

in product coordinates (x,y) (mod 1) which has no periodic solutions.

For $d\geqq 1$ use product coordinates (x,y) (mod 1) and consider the differential system

$$v_d)\qquad \begin{aligned}\dot{x}&=\cos(2\pi d\, y)\\ \dot{y}&=\sin(2\pi d\, y).\end{aligned}$$

An easy sketch of the solution curves shows that v_d lies in $\mathcal{V}_0(d,0)$ and has exactly $2d$ periodic solutions. Q. E. D.

COROLLARY. *Let $v\in\mathcal{V}_0(T^2)$. If no solution of v is periodic or if every solution of v is periodic, then v has degree $(0,0)$.*

Proof. If v has no periodic solution then we obtain a C^∞ simple closed curve $\mathcal{B}$ nowhere tangent to v, as in the theorem. Then it is easy to compute the degree $(0,0)$ for v, as is done in the theorem.

If every solution of v is periodic, say (without loss of generality) homotopic to the longitude circle $y=0$, then $v\in\mathcal{V}_0(d,0)$ for some $d\geqq 0$. But here we can choose the transversal curve $\mathcal{B}$ so that every solution of v meets $\mathcal{B}$ in just one point and is not tangent to $\mathcal{B}$. This can be done by using the existence of local cross-sections for v and also the fact that the integral curves of v are cyclically ordered by the regions into which they separate T^2. Then the degree of v around the circle $y=0$ and also around the curve $\mathcal{B}$ is zero. Thus the degree of v is $(0,0)$. Q. E. D.

THEOREM 4. *In each $\mathcal{V}_0(d,0)$ for $d\geqq 0$, the cardinality of the number of integral equivalence classes is that of the continuum.*

Proof. For $d=0$ use the differential systems

$$\dot{x}=1, \qquad \dot{y}=\lambda$$

where $\lambda > 0$ is irrational. The classical rotation number of this system on the torus (x, y) (mod 1) is λ, see [2]. Now a homeomorphism of T^2 onto itself can modify the rotation number only by an integral unimodular transformation

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad : \lambda \rightarrow \frac{a\lambda + b}{c\lambda + d} \; .$$

Thus we obtain systems which are not integrally equivalent by choosing irrational rotation numbers which are not unimodularly equivalent. Since there is only a countable set of irrational numbers unimodularly equivalent to a given λ, there must exist a continuum of such equivalence classes to constitute the continuum of all positive irrational numbers.

Now consider $d \geqq 1$ and

$$\begin{aligned} v_d \colon \dot{x} &= \cos^3(2\pi d\, y) \\ \dot{y} &= \sin^3(2\pi d\, y), \quad 0 \leqq x, y \leqq 1, \end{aligned}$$

but take the torus to be x (mod 1) and y (mod 2). On $0 \leqq x \leqq 1$ and $1 \leqq y \leqq 2$ we supplement v_d by another differential system v_s defined below, so that the entire system is in C^1 on T^2 and still has degree $(d, 0)$. Let K be a compact subset of the real interval $1 \leqq y \leqq 2$ defined as follows:

$$K = K_0 \cup K_1 \cup K_2 \cup \cdots \cup K_n \cup \cdots \cup (2).$$

Here K_0 is the interval $1 \leqq y \leqq 5/4$,

K_1 is the interval $3/2 \leqq y \leqq 3/2 + 1/8$

K_2 is the interval $7/4 \leqq y \leqq 7/4 + 1/64$

.

.

.

K_n is the interval $\dfrac{2^{n+1} - 1}{2^n} \leqq y \leqq \dfrac{2^{n+1} - 1}{2^n} + \dfrac{1}{2^{3n}}$.

Let s be a real number on $0 < s < 1$ with decimal expansion

$$s = .\, s_1 s_2 s_3 \cdots$$

where each integer s_n is on $0 \leqq s_n \leqq 9$, and say for definiteness that there is no terminating sequence of 9's. Let $K(s)$ be the compact subset of $1 \leqq y \leqq 2$ obtained by adding to K exactly s_1 points in the interval between K_1 and K_2, say equally spaced, then s_2 points between K_2 and K_3, and s_n points equally spaced in the open interval between K_n and K_{n+1}, for $n = 1, 2, 3, \cdots$. Then it is clear that if s and s' are distinct real numbers, $0 < s, s' < 1$, there is no homeomorphism of the interval $1 \leqq y \leqq 2$ onto itself which maps $K(s)$ onto $K(s')$.

For each $0 < s < 1$ define v_s in $0 \leqq x \leqq 1$, $1 \leqq y \leqq 2$ in T^2 by

$$v_s: \dot{x} = 1$$
$$\dot{y} = \phi_s(y).$$

Here $\phi_s(y) \geqq 0$ is C^∞ on $1 \leqq y \leqq 2$ and is zero just for $s \in K(s)$. We need only define $\psi_s(y) > 0$ in each open interval of $[1 \leqq y \leqq 2] - K(s)$ so that $\psi_s(y)$ is in C^∞ in the closure of this interval and has all derivatives zero at the endpoints. Such a construction is standard, see [12], and then define

$$\phi_s(y) = \psi_s(y)\, e^{1/(y-2)}$$

to obtain the required differentiability at $y = 2$. Now v_s has periodic solutions $x = t$, $y = y_0$ at just the values $y_0 \in K(s)$. In between the periodic solutions of v_s are spirals approaching these periodic solutions.

The entire differential system on T^2 consisting of v_d on $0 \leqq y \leqq 1$ and v_s on $1 \leqq y \leqq 2$ forms a C^1 system in $\mathfrak{V}_0(d, 0)$. It is easy to see that for real $s \neq s'$ the corresponding differential systems on T^2 are not integrally equivalent.[6] Q. E. D.

THEOREM 5. *Let $N \geqq 1$ be an integer and consider all differential systems $v \in \mathfrak{V}_0(T^2)$ each of which has exactly N periodic solutions. Then there are at most 4^N integral equivalence classes of such differential systems in $\mathfrak{V}_0(T^2)$ having precisely N periodic solutions.*

Proof. Let $v \in \mathfrak{V}_0(T^2)$ have exactly $N \geqq 1$ periodic solutions. We shall show that v is integrally equivalent to one of at most 4^N standard models, cf. [5, 10].

By using a preliminary diffeomorphism of T^2 onto $S^1 \times S^1$, we can assume that each periodic solution of v lies in the homology class of the cycle $y = 0$ in the product coordiates $(x, y) \pmod 1$ in $S^1 \times S^1$. Next we show that we can further assume that each periodic solution lies on a horizontal line $y =$ constant.

To see this consider the closed regions $R_1, R_2, \cdots, R_N$ in T^2 bounded by the N periodic solutions of v. Each R_i is diffeomorphic with a plane annulus A_i bounded by nonsingular C^2 simple closed Jordan curves. By results of conformal mapping theory [13], each R_i is C^1-homeomorphic to a standard circular annulus A_i. Thus the entire torus T^2 is the union of the plane annuli, with certain identifications prescribed by C^1-homeomorphisms of the bounding circles. Further each A_i is C^1-homeomorphic to a plane square Q_i: $0 \leqq x, y \leqq 1$ with the standard identification $(0, y) \cong (1, y)$ and we can also require that the induced homeomorphisms between the sides $y = 0$ or 1

[6] Putting $d = 0$ in this construction—or considering just the fields v_s for $0 < s < 1$—yields again a continuum of different integral equivalence classes in $\mathfrak{V}_0(0, 0)$.

for consecutive squares Q_i be obtained by identifying appropriate points having the same x-coordinates. This apparent simplification of the identifications on the squares Q_i can be achieved since any orientation-preserving C^1-homeomorphism of S^1 onto itself is C^1-isotopic to the identity map (as is easily proved by using elementary convex combinations of the mapping functions) and so the map between the boundaries of A_i and Q_i can be extended to the required map of A_i onto Q_i. This defines a homeomorphism h of T^2 onto $\cup Q_i$, the joined Q_i with identifications as described above. Thus h is a homeomorphism of T^2 onto T^2 which carries v onto $\bar{v}$. The vector field $\bar{v}$ on T^2 has N periodic solutions, each of which lies along a line $y=$ constant, in appropriate product coordinates. We note that $\bar{v}$ in each square Q_i, that is in the closed annular region bounded by periodic solutions on T^2, is in class C^1. We can also take $\bar{v}$ to be a unit vector field in the product coordinates in each Q_i.

We enumerate the possible integral equivalence classes of v by showing that $\bar{v}$ in each Q_i is integrally equivalent to one of four standard models in the square $Q: 0 \leqq x, y \leqq 1$,

$$\begin{array}{lll} \text{a)} & \dot{x}=1, & \dot{y}=y(1-y) \text{ or } v_a \\ \text{b)} & \dot{x}=1, & \dot{y}=y(y-1) \text{ or } v_b \\ \text{c)} & \dot{x}=\cos \pi y, & \dot{y}=y(1-y) \text{ or } v_c \\ \text{d)} & \dot{x}=\cos \pi y, & \dot{y}=y(y-1) \text{ or } v_d. \end{array}$$

Note that the additional four models obtained from these by replacing t by $-t$ need not be counted. For we can specify $\dot{x}>0$ on a selected periodic solution of $\bar{v}$ and then a choice of just one of four models is possible for the adjoining annular region on T^2. Once this choice is made there is again a choice of just one of four models (either those above or those obtained by $t \rightarrow -t$) for the next annular region, in an obvious cyclic ordering. Thus for each annular region between periodic solutions of $\bar{v}$ on T^2 there will be a choice of just one out of four cases. So we obtain an upper bound of 4^N integral equivalence classes for v on T^2, as stated in the theorem.[7]

By the Poincaré-Bendixson theory of differential equations in a plane annulus, we find that $\bar{v}$ in Q_i must belong to exactly one of the four cases:

1) $\dot{x}=1$ on $y=0$, $\dot{x}=1$ on $y=1$
each solution of $\bar{v}$ has $(y=0)$ as $(-)$ limit set and
$(y=1)$ as $(+)$ limit set

[7] A differential system on T^2 composed of standard models a), b), c), d) in consecutive annuli is in general not of class C^1, but it is of sufficient regularity (Lipschitz continuous) to insure uniqueness of solutions, and this is sufficient for the study of integral equivalence relations as it is done here.

2) $\dot{x}=1$ on $y=0$, $\dot{x}=1$ on $y=1$
each solution of $\bar{v}$ has $(y=1)$ as $(-)$ limit set and $(y=0)$ as $(+)$ limit set

3) $\dot{x}=1$ on $y=0$, $\dot{x}=-1$ on $y=1$
each solution of $\bar{v}$ has $(y=0)$ as $(-)$ limit set and $(y=1)$ as $(+)$ limit set

4) $\dot{x}=1$ on $y=0$, $\dot{x}=-1$ on $y=1$
each solution of $\bar{v}$ has $(y=1)$ as $(-)$ limit set and $(y=0)$ as $(+)$ limit set.

We shall prove that in each case 1), 2), 3), or 4) $\bar{v}$ in Q_i is equivalent to the model a), b), c), or d), respectively, in Q.

Consider case 1). We shall construct a transversal $\Gamma(s)$, a nonsingular C^1-homeomorph of $0 \leqq s \leqq 1$ in Q_i, joining the point $x=\frac{1}{2}$, $y=0$ to $x=\frac{1}{2}$, $y=1$ and such that $\Gamma(s)$ is nowhere tangent to $\bar{v}$. To construct $\Gamma(s)$ use short segments of the line $x=\frac{1}{2}$ near $y=0$ and near $y=1$, and then join these segments by a smooth curve which lies close to a solution of $\bar{v}$, but so that $\Gamma(s)$ crosses the solution curves of $\bar{v}$ and is nowhere tangent to $\bar{v}$. Of course, $\Gamma(s)$ might wind many times around the annulus Q_i (we use Q_i to designate the square or the annulus obtained by the identification $(0,y)\cong(1,y)$), but this is not significant.

Let $\Gamma_a(\sigma)$ be the transversal $x=\frac{1}{2}$, $y=\sigma$ on $0\leqq\sigma\leqq 1$ for the standard model v_a in Q. The solutions of $\bar{v}$ in Q_i define a homeomorphism ϕ of $\Gamma(s)$ onto itself: follow the solution of $\bar{v}$ initiating at each point of $\Gamma(s)$ around to the next intersection with $\Gamma(s)$. A similar homeomorphism ϕ_a is defined for $\Gamma_a(\sigma)$. Note that both ϕ and ϕ_a define homeomorphisms of the compact segment $[0,1]$ onto itself such that the endpoints are fixed, but no other points are fixed. But it is easy to see that ϕ and ϕ_a define conjugate homeomorphisms of $[0,1]$ (for note that the coordinate change $\tau=\ln\dfrac{s}{1-s}$ makes ϕ on $0<s<1$ conjugate to an orientation-preserving, fixed-point free homeomorphism of R^1 onto itself and so conjugate to a translation map $\tau\to\tau+1$). Therefore we re-parametrize $\Gamma(u)$ and $\Gamma_a(u)$ on $0\leqq u\leqq 1$ so that the homeomorphisms ϕ and ϕ_a are precisely the same, as defined on the interval $0\leqq u\leqq 1$.

Now we extend the homeomorphism

$$\Gamma(u)\to\Gamma_a(u)$$

to a homeomorphism of Q_i onto Q so as to carry the integral curves of $\bar{v}$ onto those of v_a, the standard model. To perform the required extension each solution of $\bar{v}$ in Q_i is mapped onto a solution of v_a having the same u-intercept

on $\Gamma(u)$ or $\Gamma_a(u)$. For each solution S of $\bar{v}$ we measure the time duration $T(u)$ between successive intercepts on $\Gamma(u)$ and similarly we compute $T_a(u)$ for the corresponding solution S_a of v_a. Then points on S are mapped to points of S_a so that they correspond to the same proportions of the time durations between successive intercepts with $\Gamma(u)$ or $\Gamma_a(u)$, respectively. This construction proves that $\bar{v}$ in Q_i as in case 1) is integrally equivalent to v_a in Q, and the homeomorphism can be taken as $x \leftrightarrow x$ on the sides $y = 0$ and $y = 1$.

For $\bar{v}$ in Q_i in case 2) we can conclude, just as for case 1), that $\bar{v}$ is integrally equivalent to v_b in Q. We next consider $\bar{v}$ in Q_i in case 3). We can assume that $\dot{x} = -1$ for y near 1 and $\dot{x} = +1$ for y near 0.

For v_c in Q we construct the transversal $\Sigma_c(\theta)$ on $0 \leqq \theta < 1$ consisting of the circle $y = \frac{1}{2}$. For $\bar{v}$ in Q_i we also construct a circular transversal $\Sigma(\theta)$. To do this we take a segment γ on the line $L: x = \frac{1}{2}$ near $y = 1$ (in the band wherein $\dot{x} = -1$) between two successive intersections of a solution S of $\bar{v}$ with L. Now we modify the upper end of γ by sliding it back along S until we form a simple closed Jordan curve $\Sigma(\theta)$, $0 \leqq \theta < 1$, which is nonsingular, differentiable, nowhere tangent to $\bar{v}$. Also $\Sigma(\theta)$ meets every solution of $\bar{v}$, other than $y = 0$ and $y = 1$, in exactly one point. We shall map $\bar{v}$ in Q_i onto v_c in Q by mapping the solution $S(\theta)$ of $\bar{v}$ onto the solution $S_c(\theta)$ of v_c with the same θ-intercepts on $\Sigma(\theta)$ and $\Sigma_c(\theta)$.

On the line L let ℓ_1 be a segment between consecutive hits of $S(\theta)$ after leaving $\Sigma(\theta)$. Assume ℓ_1 lies within the band wherein $\dot{x} = -1$ for $\bar{v}$. Every other solution of $\bar{v}$ meets ℓ_1 in just one point. Map the arc of $S(\theta)$ between the intercept with $\Sigma(\theta)$ and the first meeting with ℓ_1 onto the corresponding arc of v_c, using proportional time durations to define the map. Map the arc of $S(\theta)$ following the first meeting with ℓ_1 onto the corresponding arc of $S_c(\theta)$ by keeping the same x-coordinates for corresponding points. Map $y = 1$ in Q_i onto $y = 1$ in Q by $x \leftrightarrow x$.

A similar construction maps the arc of $S(\theta)$ preceding the intercept with $\Sigma(\theta)$ onto the corresponding arc of $S_c(\theta)$. We perform this map for each $S(\theta)$ onto $S_c(\theta)$ for $0 \leqq \theta < 1$. Thus we obtain the required homeomorphism carrying $\bar{v}$ in Q_i onto v_c in Q.

For $\bar{v}$ in Q_i in case 4) we can conclude, just as in case 3), that $\bar{v}$ is integrally equivalent to v_d in Q.

We finally conclude that v on T^2 is integrally equivalent to a C^1 vector field $\bar{v}$ on T^2. Each of the N periodic solutions of $\bar{v}$ is along a longitude circle $y = \text{constant}$, and each of the annular regions bounded by a pair (or just one if $N = 1$) of periodic solutions is filled by models a), b), c), or d) (or a model with $t \to -t$). These canonical forms for v on T^2 show that there are at most 4^N integral equivalence classes for v. Q. E. D.

It is easy to show that the bound 4^N is not best possible, in fact $2 \cdot 4^{N-1}$ is a bound as is seen immediately, but it seems difficult to find the best bound. Using Theorem 5 we can obtain special results for real analytic differential systems $v \in \mathfrak{V}_0(T^2)$ on the real analytic torus T^2. Of course the set $\mathfrak{V}_0(0,0)$ contains a continuum of non-equivalent analytic vector fields. But for $\mathfrak{V}_0(d,0)$ with $d > 0$, the following results are of interest.

LEMMA. *A real analytic vector field $v \in \mathfrak{V}_0(T^2)$ on a real analytic torus T^2 either has a finite number of periodic solutions or else every solution is periodic.*

Proof. If there were an infinite number of periodic solutions, then we investigate the behavior of v in the neighborhood N of a limit periodic solution. An elementary consequence of analyticity is that every solution of v in N is periodic. Examine the boundary solutions of N and note that each of these lies interior to a neighborhood filled with periodic solutions of v. It thus follows that all T^2 is filled by periodic solutions of v. Q. E. D.

THEOREM 6. *On the real analytic torus $T^2 = S^1 \times S^1$, in each $\mathfrak{V}_0(d,0)$, for $d > 0$, there exists precisely a countable infinity of integral equivalence classes which contain analytic vector fields.*

Proof. Let $v \in \mathfrak{V}_0(d,0)$ be analytic on T^2. Then v has just a finite number $N \geqq 2d$ of periodic solutions and so v belongs to one of 4^N (or less) integral equivalence classes of $\mathfrak{V}_0(T^2)$. By counting the equivalence classes corresponding to $N = 1, 2, 3, \cdots$ periodic solutions, we obtain a countable number of integral equivalence classes represented by analytic vector fields in $\mathfrak{V}_0(d,0)$. Q. E. D.

For structurally stable differential systems on the torus T^2 the possible integral equivalence classes are rather simple. Note that a structurally stable system is always integrally equivalent to a real analytic system.

THEOREM 7. *A structurally stable differential system $v \in \mathfrak{V}_0(T^2)$ has a finite number $N \geqq 1$ of periodic solutions.*

Proof. If v has no periodic solutions, then v is integrally equivalent to a linear system

$$\dot{x} = 1, \qquad \dot{y} = \lambda \qquad (x \text{ and } y \bmod 1)$$

with irrational λ. However, even using a C^1-diffeomorphism of T^2 onto itself, we can map v onto a differential system

$$v' : \dot{x} = 1, \ \dot{y} = h(x,y)$$

with $h(x,y)$ in C^1 on T^2. The system v' must be structurally stable and yield a rotation number of λ, that is, the map of the circle $x = 0$ of T^2 onto itself

defined by the solution curves of v' is conjugate to the irrational rotation of S^1 through an angle of $2\pi\lambda$.

But for arbitrarily small $\epsilon > 0$ the perturbed system

$$v'_\epsilon: \dot{x} = 1,\ \dot{y} = h(x, y) + \epsilon$$

has a rational rotation number and hence a periodic solution. Thus v'_ϵ is not integrally equivalent to v', which contradicts the hypothesis of structural stability of v. Therefore v has at least one periodic solution.

Since v is integrally equivalent to a real analytic system on T^2, the Lemma together with a consideration as above shows that v has only a finite number of periodic solutions on T^2. Q. E. D.

COROLLARY. *There is just a countable number of integral equivalence classes which contain structurally stable differential systems in* $\mathfrak{V}_0(T^2)$.

It is easy to see that each $\mathfrak{V}_0(a, b)$ contains a countable number of inequivalent structurally stable equations. One should note, however, that $v_1 \in \mathfrak{V}_0(T^2)$ can be integrally equivalent to a structurally stable system $v \in \mathfrak{V}_0(T^2)$ and yet v_1 can fail to be structurally stable. This is possible if we recall that a periodic solution of a structurally stable system v must have a non-zero characteristic exponent. Then v_1 can have the same topological properties as v and yet have periodic solutions with zero characteristic exponents, cf. [7, 8], and see Footnote 3.

M. Peixoto has proved that the structurally stable differential systems in $\mathfrak{V}(T^2)$ are dense in $\mathfrak{V}(T^2)$, see [9].

We now turn to the problem of bifurcation of differential systems on T^2. If v_0 and v_1 are in $\mathfrak{V}_0(T^2)$, can they be joined by a continuous family of differential systems $v_\lambda \in \mathfrak{V}_0(T^2)$ on $I: 0 \leqq \lambda \leqq 1$ with only a finite number of bifurcation values of $\lambda \in I$? Of course we assume v_0 and v_1 lie in the same component of $\mathfrak{V}_0(T^2)$, and it is no restriction to take this component as $\mathfrak{V}_0(d, 0)$.

THEOREM 8. *On the torus* $T^2 = S^1 \times S^1$ *let* v_0 *and* $v_1 \in \mathfrak{V}_0(d, 0)$ *for* $d > 0$. *Then there exists a continuous family of differential systems*

$$v_\lambda \in \mathfrak{V}_0(d, 0) \quad for \quad \lambda \in I: 0 \leqq \lambda \leqq 1,$$

joining v_0 *to* v_1, *and such that* v_λ *has only a finite set of bifurcation values in* I.

Proof. Using the techniques employed in the proof of Theorem 5 we can decompose the integral curve family of v_0 on T^2 into a closed set filled by periodic solutions and a countable number of annuli Q_i, each integrally equivalent to one of the four standard models a), b), c), d) (and these with

$t \to -t$), as discussed above. Since v_0 has no critical point on T^2, only a finite number of the "horseshoe" annuli of types c) and d) can occur. The union of the finitely many closed annuli of types c) and d) is a closed set in T^2 and its complement is open in T^2; it can be described by an open set on a meridian circle S^1, that is, by a finite union of disjoint intervals $I_1, I_2, \cdots, I_N$ in S^1.

Now define v_λ on $0 \leqq \lambda \leqq \frac{1}{4}$ to be v_0 for $\lambda = 0$ and for $\lambda > 0$ define v_λ by the integral curve family obtained from v_0 by narrowing each of the open intervals $I_1, I_2, \cdots, I_N$ by linear contraction. At $\lambda = \frac{1}{4}$ each of the intervals $I_1, \cdots, I_N$ is replaced by a single point which corresponds to a periodic solution of $v_{1/4}$. Then $\lambda = \frac{1}{4}$ is the first bifurcation value of $\lambda \in I$, and $v_{1/4}$ possesses only a finite number of periodic solutions. Also $v_{1/4}$ has only a finite number of "horseshoe" annuli and if these are all of the type c) (or c_) obtained by replacing t by $-t$), there are exactly $2d$ such annuli. Since $d > 0$, the vector field $v_{1/4}$ cannot consist entirely of annuli of types d) and d_).

We next show that if there is an annulus of type c) and also one of type d) in $v_{1/4}$, then a continuous deformation v_λ with two bifurcations can reduce the number of d) annuli by one. For consider adjacent annuli of types c) and d), say,

$$\dot{x} = \cos \pi y \qquad 0 \leqq x \leqq 1$$
$$\dot{y} = y(y-1)^2(y-2) \qquad 0 \leqq y \leqq 2.$$

Now define the parametrized family on the two annuli, $0 \leqq x \leqq 1$, $0 \leqq y \leqq 2$, by

$$\dot{x} = \cos \pi y$$
$$\dot{y} = y(y-1)^2(y-2) + y(y-2)(\lambda - \tfrac{1}{4}) \quad \text{for } \lambda \geqq \tfrac{1}{4}.$$

But for $\lambda > \frac{1}{4}$ this differential system is integrally equivalent to an annular region of type b). Then again we deform v_λ to contract this b) annulus to a single periodic solution curve. Thus with two bifurcation values, we can extend the definition of v_λ, say to $0 \leqq \lambda \leqq \frac{3}{8}$, to obtain a system with one less d)-annular region than occurs in $v_{1/4}$.

In this way we can define v_λ on $0 \leqq \lambda \leqq \frac{1}{2}$ so that $v_{1/2}$ has an integral curve family composed of just $2d$ annuli of type c) and c_), which occur alternately in the cyclic ordering of the annuli around T^2. In the same way we can define v_λ on $\frac{1}{2} \leqq \lambda \leqq 1$ which joins v_1 to the same standard model $v_{1/2}$. Thus the total parametrized family v_λ on $0 \leqq \lambda \leqq 1$ has only a finite number of bifurcation values. Q. E. D.

The problem of bifurcation for a continuous family v_λ joining v_0 to v_1

in $\mathfrak{V}_0(0,0)$ is more difficult and will not be analysed completely here. Of course if v_λ is required to remain within the open subset Q of $\mathfrak{V}_0(0,0)$ for which the classical rotation number is well-defined, then the bifurcation analysis is easy. For instance, suppose v_0 and v_1 in $Q \subset \mathfrak{V}_0(0,0)$ have different rotation numbers and $v_\lambda \in Q$ for $0 \leqq \lambda \leqq 1$. Since the rotation number of v_λ is $\rho(\lambda)$, which is continuous in λ, we find that $\rho(\lambda)$ must take on a non-countable number of modularly inequivalent irrational values. But this means that v_λ has a continuum of bifurcation values.

UNIVERSITY OF MINNESOTA.

REFERENCES.

[1] L. Ahlfors and L. Sario, *Riemann Surfaces*, Princeton (1960).
[2] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, New York (1955).
[3] G. Hardy and E. Wright, *An Introduction to the Theory of Numbers*, Oxford (1938).
[4] B. V. Kerekjarto, *Vorlesungen über Topologie*, Berlin (1923).
[5] H. Kneser, "Kurvenscharen auf den Ringflächen," *Mathematische Annalen*, vol. 91 (1924), pp. 133-154.
[6] L. Markus, "Global structure of ordinary differential equations in the plane," *Transactions of the American Mathematical Society*, vol. 76 (1954), pp. 127-148.
[7] ———, "The behavior of the solutions of a differential system near a periodic solution," *Annals of Mathematics*, vol. 72 (1960), pp. 245-266.
[8] ———, "Structurally stable differential systems," *Annals of Mathematics*, vol. 73 (1961), pp. 1-19.
[9] M. M. Peixoto, "Structural stability on two-dimensional manifolds," *Instituto de Mathematica Pura e Aplicado*, Rio de Janeiro (1961).
[10] L. B. Reinhart, "Line elements on the torus," *American Journal of Mathematics*, vol. 81 (1959), pp. 617-631.
[11] S. Schwartzman, "Asymptotic cycles," *Annals of Mathematics*, vol. 66 (1957), pp. 270-284.
[12] N. Steenrod, *The Topology of Fibre Bundles*, Princeton (1951).
[13] S. Warschawski, "Über das Randverhalten der Ableitung der Abbildungsfunktion bei konformer Abbildung," *Mathematische Zeitschrift*, vol. 35 (1932), pp. 321-456.

ON σ-DISCRETENESS AND BOREL ISOMORPHISM.*

By A. H. Stone.[1]

1. Introduction. Throughout this paper, all spaces considered are to be metric. In [4] a start was made on the problem of classifying and characterizing the absolute Borel spaces under Borel isomorphism or generalized homeomorphism.[2] The main difficulty remaining was to find topological properties which are invariant under Borel isomorphism. It was shown that the weight of a space is one invariant; it is natural to conjecture that the local weight might be another, but this is not the case. In fact, an absolute Borel space can have arbitrarily high local weight and yet be Borel isomorphic to a discrete space (of local weight 1).[3] However, we shall prove in the present paper that it will then have to be σ-discrete;[4] in fact, σ-discreteness is invariant under Borel isomorphism (Theorem 3, 3.4). This enables us to improve (Theorem 2, 3.3) the description of the hereditarily Borel sets given in [4, 6.2], and to complete (Theorem 6, 6.1) the Borel classification of the absolute Borel spaces of weight $\aleph_1$, started in [4, 5.5]. This last result depends on an auxilliary result which may be of independent interest. One well-known application of the axiom of choice is to assign to each countable limit ordinal a sequence of smaller ordinals converging to it; and such an assignment is itself a sufficient substitute for the axiom of choice in many applications. Thus it can be expected that there is no "natural" way of assigning such sequences of ordinals, and we confirm this by proving (Theorem 5 and Corollary, 5.4 and 5.5) that no such choice of sequences can yield a Borel set in the space of all sequences of countable ordinals, in either the metric or the usual topology.[5] We also deduce from Theorem 3 a theorem about continuous Borel isomorphisms of complete spaces (Theorem 4, 4.1). This will later permit a generalization of the theory to absolute Borel spaces which are σ-locally of weight $< k$, k being an arbitrary infinite cardinal. It turns

* Received July 22, 1963.

[1] This work was supported by the National Science Foundation Grant NSF-G 19883.

[2] A "Borel isomorphism is a 1-1 mapping f such that both f and f^{-1} take Borel sets into Borel sets. If further f and f^{-1} alter Borel classes by at most fixed amounts, f is a "generalized homeomorphism."

[3] See 2.4 below for an example.

[4] See 2.1 for the definition.

[5] This is the only occurrence of non-metrizable spaces in this paper.

out that this property is also Borel invariant, and that the classification problem of [4] can thereby be solved; the author hopes to deal with this in a subsequent paper. The σ-discrete case considered here is merely the special case $k=\aleph_0$, but has to be dealt with first, as we need Theorem 4 for treating the general case.

The terminology follows [1], in general; further terminology is as in [4], acquaintance with which is presupposed. In particular, the letter k is used for an infinite cardinal. The "Baire space" $B(k)$ is the product of $\aleph_0$ discrete sets each of cardinal k; $B(\aleph_0)$ is the space of irrational numbers. The Cantor set is denoted by $C(\aleph_0)$. As stated above, all spaces considered are to be metrizable; we use the symbol ρ for a metric on the space X, and $S(x,\epsilon)$ for the ϵ-neighborhood $\{y \mid \rho(x,y)<\epsilon\}$ $(\epsilon>0)$ of x in X. The diameter of a set E is written $\delta(E)$.

2. Lemmas on σ-discreteness.

2.1. We say (as in [3]) that a subset A of (metric) space X is *relatively discrete* if each point of A is isolated in A, *discrete* (in X) if A is closed in X and relatively discrete, and *metrically discrete* (with respect to a metric ρ on X) if it is ϵ-discrete for some $\epsilon>0$; that is, if $x,y\in A$ and $x\neq y$ imply $\rho(x,y)\geqq\epsilon$. Each of these three properties may reasonably be used to define "σ-discrete"; fortunately it does not matter which we use, in view of the following lemma.

LEMMA 1. *The following statements about a subset A of a space X are equivalent:*

(1) $A=\cup\{A_n \mid n=1,2,\cdots\}$ *where each* A_n *is relatively discrete,*

(2) $A=\cup\{B_n \mid n=1,2,\cdots\}$ *where each* B_n *is discrete,*

(3) $A=\cup\{C_n \mid n=1,2,\cdots\}$ *where each* C_n *is metrically discrete.*

Since metrically discrete $\Rightarrow$ discrete $\Rightarrow$ relatively discrete, it suffices to prove (1) $\Rightarrow$ (3). From [3, p. 99], each $A_n=\cup\{C_{nm} \mid m=1,2,\cdots\}$ where C_{nm} is metrically discrete; we merely renumber $\{C_{nm}\}$ into a single sequence $\{C_n\}$.

The first of these three properties depends only on A, and is a topological invariant of A. Thus a space A which has these properties for any one containing space X (e.g., itself) has them for all X; we call A "*σ-discrete.*"

2.2. Clearly if A is σ-discrete, then so is every subset of A. We say that a space X is *locally σ-discrete* if each point of X has a σ-discrete neighborhood (which may be supposed to be open) in X.

LEMMA 2. *If X is locally σ-discrete, then X is σ-discrete.*

This follows from a theorem of Michael [2, Th. 3.6(a)]. A direct proof, by using a "σ-discrete" open refinement of an open covering of X by σ-discrete sets, is also straightforward.

2.3. LEMMA 3. *If f is a 1-1 continuous map of a space X onto a space Y, and $D \subset Y$ is σ-discrete, then so is $f^{-1}(D)$.*

It is enough to prove that if Y is discrete then so is X. But, for each $x \in X$, $f(x)$ is open in Y; hence $\{x\} = f^{-1}\{f(x)\}$ is open in X.

2.4. *An example.* Before proving, in the next section, that σ-discreteness is invariant under Borel isomorphism, we give a simple example to show that discreteness is emphatically not invariant. Consider $B(k) = \prod\{T_n \mid n = 1, 2, \cdots\}$, where T_n is a discrete space of cardinal k. For each n, pick a fixed element $a_n \in T_n$, and put E_m = set of all points $x = (x_1, x_2, \cdots)$ of $B(k)$ for which $x_n = a_n$ for all $n > m$, $m = 1, 2, \cdots$. Write $E = \cup \{E_m \mid m = 1, 2, \cdots\}$. Each E_m is metrically discrete, in the usual metric on $B(k)$,[6] so every subset of E_m is locally compact. Thus every subset of E is an absolute F_σ (see [5]; alternatively this is an easy consequence of Lemma 1 above). Hence, from [4, Th. 12, 6.2], E is Borel isomorphic (by a generalized homeomorphism of class $(0,1)$) to a discrete set of cardinal k (again this is also easy directly). But each non-empty open subset of E is arbitrarily far from being discrete.

2.5. One final remark: it is easy to see (from Lemma 1) that a σ-discrete space of weight k has exactly k points. In particular, for *separable* spaces, "σ-discrete" $\Longleftrightarrow$ "countable."

3. Invariance of σ-discreteness.

3.1. The main step in the proof is the following lemma, which combines a previous theorem with a very special case of the theorem to be proved.

LEMMA 4. *Let Y be an absolute Borel set of (additive or multiplicative) class $\alpha \geqq 1$, and of weight $\leqq k$. There exists a generalized homeomorphism f, of class $(0, \omega_0{}^\alpha)$, of a closed subset A of $B(k)$ onto Y, such that*

(*) *for each σ-discrete $E \subset A$, $f(E)$ is also σ-discrete.*

Except for the last assertion (*), this has been proved in [4, Th. 4, 3.2].

[6] Compare 5.2 below.

We shall follow through the argument there and verify that it also gives (*). It is enough, in view of Lemma 1 (2.1) to prove

(**) *If E is metrically discrete in A, then $f(E)$ is σ-discrete.*

The argument in [4, § 3] constructs A and f by transfinite induction over α, starting with the G_δ case (multiplicative class $\alpha = 1$). In this case (see [4, Lemma 3.3 and corollary]) f was shown to be of class $(0,1)$; so, since E is closed in A, $f(E)$ is G_δ in Y, and so is an absolute G_δ. Thus we may give $f(E)$ a complete metric. Let K be the perfect kernel (largest dense-in-itself subset) of $f(E)$; if $K \neq \emptyset$, K would contain a homeomorph of the Cantor set $C(\aleph_0)$ (see e.g. [1, p. 351]), which in turn would contain a non-Borel subset H; but then $f^{-1}(H)$ would be a non-Borel subset of E—which is absurd, since every subset of E is closed. So $K = \emptyset$, and therefore [6, Th. 4] $f(E)$ is σ-discrete.

The F_σ case is, for present purposes, covered by the $G_{\delta\sigma}$ case, so we proceed to the two remaining steps.

(i) Y of additive class $\alpha > 1$. Here $Y = \cup_n F_n$ $(n = 1, 2, \cdots)$ where F_n is of multiplicative class $< \alpha$, and the sets $F_1, F_2, \cdots$ are (pairwise) disjoint; and the hypothesis of induction gives $f_n(A_n) = F_n$ where f_n, A_n, satisfy the requirements of the present lemma. The construction in [4] makes $A = \cup_n (A_n \times n)$, and defines f so that $f(a_n, n) = f_n(a_n)$ $(a_n \in A_n)$. If E is discrete in A, we clearly have $E = \cup_n (E_n \times n)$ where E_n is discrete in A_n; hence $f(E) = \cup_n f_n(E_n)$ where each $f_n(E_n)$ is σ-discrete. This shows that $f(E)$ is σ-discrete, verifying (**).

(ii) Y of multiplicative class $\alpha > 1$. Here $Y = \cap_n F_n$ where F_n is of additive class $< \alpha$ and where $F_1 \supset F_2 \supset \cdots$. The induction hypothesis gives $F_n = f_n(A_n)$ as in the present lemma, where A_n is a closed subset of a copy B_n of $B(K)$. The construction in [4] defines A to be the subset of $\prod_n B_n$ consisting of all points $a = \{a_n\}$ for which $a_n \in A_n$ and $f_n(a_n) = f_1(a_1)$ $(n = 1, 2, \cdots)$; and f is defined by: $f(a) = f_1(a_1)$. We must verify that (**) holds. In doing so, we may suppose that each A_n has a metric $\rho_n \leqq 2^{-n}$, and that A (a subspace of $\prod_n A_n$) is metrized by ρ, where

$$\rho(a, b) = \sum_{n=1}^{\infty} \rho_n(a_n, b_n). \tag{1}$$

We introduce the following notation. For each positive integer m, the product $A_1 \times A_2 \times \cdots \times A_m$ is written P^m; the infinite product $\prod_n A_n$ is denoted by P. The projection of P onto A_n is π_n; the projection of P onto

P^m (that is, $\pi_1 \times \pi_2 \times \cdots \times \pi_m$) is written π^m. Observe that these projections, restricted to A, are all 1-1, because each $a \in A$ satisfies $a_r = f_r^{-1} f_s(a_s)$ $(r, s = 1, 2, \cdots)$. We metrize P^m by ρ^m, where

$$\rho^m(a, b) = \sum_{n=1}^{m} \rho_n(a_n, b_n). \tag{2}$$

Now suppose $E \subset A$ is ϵ-discrete, $\epsilon > 0$. Then, if m is large enough, we clearly have from (1) and (2) that

(3) $\quad \pi^m(E)$ is metrically discrete in P^m.

We shall show that (3) implies

(4) $\quad \pi_1(E)$ is σ-discrete.

To do this, we use induction over m. The case $m = 1$ being trivial, assume (3) $\Rightarrow$ (4) for a particular m, and for all $E \subset A$, and suppose that $\pi^{m+1}(E)$ is a δ-discrete subset of $P^{n+1} = P^n \times A_{n+1}$. Write $E^{m+1} = \pi^{m+1}(E)$, $E^m = \pi^m(E) =$ projection of E^{m+1} on P^m. Let U be any open subset of P^m of ρ^m-diameter $< \delta/2$. We shall show that

(5) $\quad U \cap E^m$ is σ-discrete,

and as the first step in establishing (5) we prove

(6) $\quad \pi_{m+1}(E^{m+1} \cap (U \times A_{m+1}))$ is discrete.

In fact, if b_{m+1} and c_{m+1} are distinct points of this set, there exist (unique) points $b, c \in E$ such that $b_{m+1} = \pi_{m+1}(b)$ and $c_{m+1} = \pi_{m+1}(c)$. Thus $\pi^{m+1}(b)$ and $\pi^{m+1}(c)$ are distinct points of $\pi^{m+1}(E)$, so by hypothesis their ρ^{m+1}-distance is $\geqq \delta$; that is, from (2), $\rho^m(\pi^m(b), \pi^m(c)) + \rho_{m+1}(b_{m+1}, c_{m+1}) \geqq \delta$. But U has diameter $< \delta/2$, so $\rho^m(\pi^m(b), \pi^m(c)) < \delta/2$. Hence $\rho_{m+1}(b_{m+1}, c_{m+1}) \geqq \delta/2$, proving (6).

Now f_{m+1} sends discrete sets into σ-discrete ones; hence $f_{m+1}\pi_{m+1}(E^{m+1} \cap (U \times A_{m+1}))$ is σ-discrete. Let f^m denote the (continuous, 1-1) map of $\pi^m(A)$ given by: $f^m(a_1, \cdots, a_m) = f_1(a_1) = \cdots = f_m(a_m)$; then $f^m \pi^m$ coincides with $f_{m+1}\pi_{m+1}$ on E^{m+1}, and so $f^m \pi^m(E^{m+1} \cap (U \times A_{m+1}))$ is σ-discrete. By Lemma 3 (2.3) it follows that $\pi^m(E^{m+1} \cap (U \times A_{m+1}))$ is σ-discrete; but this set is just $E^m \cap U$, so (5) is proved.

From (5), E^m is locally σ-discrete; Lemma 2 (2.2) now shows that E^m is σ-discrete and thus the union of a sequence of metrically discrete sets $D^m(i)$, $i = 1, 2, \cdots$. We can write $E = \cup_i D(i)$ where $\pi^m(D(i)) = D^m(i)$.

From our induction hypothesis it follows that each $\pi_1(D(i))$ is σ-discrete; hence so is their union, $\pi_1(E)$.

This proves, then, that (3) implies (4). Since (3) is true for large enough m, (4) must be true; but f_1 preserves σ-discreteness, so finally $f(E) = f_1\pi_1(E)$ is σ-discrete, completing the proof of the lemma.

3.2. THEOREM 1. *For each absolute Borel space Y, one and only one of the following alternatives is true: either* (i) *Y is σ-discrete, or* (ii) *Y contains a subset homeomorphic to the Cantor set $C(\aleph_0)$.*[7]

Proof. Apply Lemma 4 to obtain $Y = f(A)$ where A is closed in $B(k)$ and f is continuous and 1-1 and preserves σ-discreteness. Consider the perfect kernel K of A. If $K = \emptyset$, A is σ-discrete [6, Th. 4], and therefore so is Y. If $K \neq \emptyset$, K contains a closed subset C homeomorphic to $C(\aleph_0)$ [1, p. 351]; since C is compact, $f \mid C$ is a homeomorphism, so that $f(C)$ is a subset of Y homeomorphic to $C(\aleph_0)$. Finally, (i) and (ii) cannot hold simultaneously, as that would imply that $C(\aleph_0)$ is σ-discrete and so (from 2.5) countable.

3.3. The preceding theorem can be used to improve a characterization [4, Th. 12, p. 26] of the absolute Borel sets which are Borel isomorphic to discrete sets. The results is:

THEOREM 2. *The following statements about an arbitrary (metric) space X are equivalent:*

(1) *Every subset of X is an absolute Borel set.*

(2) *Every subset of X is an absolute F_σ.*

(3) *Every G_δ subset of X is an absolute F_σ.*

(4) *X is σ-discrete.*

(5) *X is absolutely Borel, and is Borel isomorphic (or generalized homeomorphic) to a discret set.*

(6) *X is absolutely Borel, and contains no homeomorph of $C(\aleph_0)$.*

(7) *X is absolutely Borel, and every separable subset of X is countable.*

(8) *X is absolutely Borel, and for every $Y \subset X$ we have $\| \bar{Y} \| = \| Y \|$.*

Proof. Since every σ-discrete space is absolutely Borel (it is an absolute F_σ, by Lemma 1), the equivalence of (4) and (6) has just been shown. The equivalence of (1), (5) and (7), and the implication (7) $\Rightarrow$ (8), were proved in [4], loc. cit. The implications (8) $\Rightarrow$ (7), (2) $\Rightarrow$ (3) and (in

[7] For separable spaces, Theorem 1 reduces to the well-known theorem (of Alexandroff and Hausdorff) that every separable Borel set Y is either countable or contains a copy of $C(\aleph_0)$; see [1, p. 355].

view of Lemma 1) $(4) \Rightarrow (2) \Rightarrow (1)$ are trivial. We must prove $(1) \Rightarrow (4)$ and $(3) \Rightarrow (4)$. Suppose that (1) or (3) holds but (4) fails. Then X is itself an absolute Borel set; hence, by Theorem 1, X contains a subset homeomorphic to $C(\aleph_0)$. But this contains a non-Borel subset, contradicting (1); and it also contains a subset homeomorphic to the irrational numbers, contradicting (3).

Remark. The generalized homeomorphism in (5) may be taken to be of class $(1,0)$. In fact, from (2), *every* 1-1 mapping of X onto a discrete set will automatically be of class $(1,0)$. It is interesting to contrast (3) with the "dual" statement that every F_σ subset of X is an absolute G_δ; it is not hard to see that this is equivalent to saying that X is scattered, or equivalently that every subset of X is absolutely both F_σ and G_δ, or that X is σ-discrete and an absolute G_δ; see [6, Th. 11] for further characterizations.

3.4. As a corollary, the equivalence of (4) and (5) gives the Borel invariance of σ-discreteness, which we state as:

THEOREM 3. *If an absolute Borel set X is Borel isomorphic to a σ-discrete space, then X is σ-discrete.*

It may be worth calling attention to another consequence of Theorem 2 (equivalence of (1), (2) and (5)): if an absolute Borel set X has a subset which is not F_σ in X, then X has Borel subsets of arbitrarily high exact class ($< \omega_1$), and also has non-Borel subsets (and in fact non-analytic subsets, as the proof of Theorem 2 shows). It would be interesting to know whether this continues to hold without the requirement that X be absolutely Borel; and the same applies to Theorem 3.

4. Continuous Borel isomorphisms.

4.1. A continuous Borel isomorphism is, of course, continuous and 1-1, and it takes every σ-discrete set onto a σ-discrete set (Theorem 3). These are the only properties really required of the mapping f in the following theorem, but we state the theorem in the more restrictive form in which it will be quoted in subsequent work.[8]

THEOREM 4. *Let k be an infinite cardinal and X a non-empty complete (metric) space in which every non-empty open set has weight k. Let f be a*

[8] Genuinely more restrictive, for it is easy to give an example of a continuous 1-1 mapping f of a complete space X such that f preserves σ-discreteness but is not a Borel isomorphism. (Roughly, take X to be the discrete union of separable spaces X_λ, $\lambda < \omega_1$; and arrange that f sends a G_δ subset of X_λ to a subset of X_λ of exact class λ.)

continuous Borel isomorphism of X onto an absolute Borel space Y. Then there exists a closed subset C^* of X such that (i) *C^* is homeomorphic to $B(k)$ (or to $C(\aleph_0)$ if $k=\aleph_0$),* (ii) *$f(C^*)$ is closed in Y,* (iii) *$f \mid C^*$ is a homeomorphism.*

Remark. This theorem resembles [4, Th. 5, 4.1]; but the hypotheses here are different, and the conclusion is stronger. In [4, Th. 5] the map f was allowed to be merely continuous; but it should be noted that here it would not suffice to have f continuous and 1-1 even if $f(X)$ is complete. This is shown by the following example. Let $f_n\colon T_n \to I_n$ $(n=1,2,\cdots)$ be, for each n, any 1-1 map of a discrete set T_n of c points onto a copy I_n of the closed unit interval. Then the product map $f=\prod_n f_n$ is a continuous 1-1 map of $\prod_n T_n = B(c)$ onto the Hilbert cube $I^{\aleph_0}$. Every non-empty open subset of $B(c)$ has weight c [4, p. 6]; but $I^{\aleph_0}$ does not contain any homeomorph of $B(c)$, since the weight of $I^{\aleph_0}$ is only $\aleph_0$.

4.2. The proof of Theorem 4 is basically similar to that of [4, Th. 5], but the details are simpler here. We start an inductive construction by picking $c_0 \in X$, and writing $U(c_0)=S(c_0,1)=\{x \mid x\in X, \rho(x,c_0)<1\}$, where ρ is the metric on X. We suppose first that k *is not sequential*; then the neighborhood $S(c_0,\frac{1}{2})$ has, for some $\eta>0$, an η-discrete subset E_1 of cardinal k (see [3, p. 100]). Now $f(E_1)$ is σ-discrete and therefore (Lemma 1) contains, for some $\eta'>0$, an η'-discrete subset D_1 of cardinal k also. Put $C_1(c_0)=f^{-1}(D_1)$. From $C_1(c_0)$ we omit any point whose distance from c_0 is less than $\eta/2$, and also any point x such that $\rho'(f(x),f(c_0))<\eta'/2$, ρ' denoting the metric on Y; there can be at most one omitted point of each kind. To save notation we refer to the remaining set as $C_1(c_0)$ also; its cardinal is still k. Taking $\epsilon_1(c_0)$ to be a small enough positive number, and writing $U(c_1)=S(c_1,\epsilon_1(c_0))$ for each $c_1\in C_1(c_0)$, we thus obtain, for all $c_1, c_1' \in C_1(c_0)$,

$$U(c_1)\subset U(c_0);\quad \delta(U(c_1))<\tfrac{1}{2};\quad \delta'(f(U(c_1)))<\tfrac{1}{2};$$

$$\text{if } c_1\neq c_1',\text{ then } \rho(U(c_1),U(c_1'))>\epsilon_1(c_0) \text{ and}$$

$$\rho'(f(U(c_1)),f(U(c_1')))>\epsilon_1(c_0);\quad \rho(c_0,U(c_1))>\epsilon_1(c_0);$$
$$\rho'(f(c_0),f(U(c_1')))>\epsilon_1(c_0);\quad \| C_1(c_0)\| = k;\quad \epsilon_1(c_0)<\tfrac{1}{2}.$$

We repeat the operation on each $U(c_1)$ $(c_1\in C_1(c_0))$, obtaining a metrically discrete set $C_2(c_1)\subset U(c_1)$ of k points, mapped by f onto a metrically discrete subset of Y, and neighborhoods $U(c_2)=S(c_2,\epsilon_2(c_1))$ $(c_2\in C_2(c_1))$ at distances greater than $\epsilon_2(c_1)$ from each other and from c_1,

where $0 < \epsilon_2(c_1) < \epsilon_1(c_0)/3$, and such that the sets $f(U(c_2))$ have similar properties. We put $C_2 = \cup \{C_2(c_1) \mid c_1 \in C_1(c_0)\}$, and repeat the argument on each $U(c_2)(c_2 \in C_2)$; and so on. This achieves substantially the same set-up as [4, p. 17]; and the rest of the proof is essentially the same as [4, p. 18]; we here summarize it briefly, referring to [4] for the details. We consider all the "descendant sequences" $c_0, c_1, \cdots, c_i, \cdots$, where $c_i \in C_i(c_{i-1})$, $i = 1, 2, \cdots$; all of them are Cauchy, and so converge; we take C^* to be the set of all limits of such sequences. It is not hard to see that C^* is closed in X, $f(C^*)$ is closed in Y, $f \mid C^*$ is a homeomorphism, and C^* is homeomorphic to $B(k)$.

If k is sequential, say $k = \sup p_n$, $p_1 < p_2 < \cdots$, we modify the preceding construction. First suppose $k > \aleph_0$; then we may assume $p_1 > \aleph_0$. We take E_1 to be of cardinal p_2, and D_1 of cardinal p_1; thus $C_1(c_0)$ will be of cardinal p_1 instead of k. Similarly we make each $C_2(c_1)$ of cardinal p_2, and so on. As in [4, p. 19], the resulting C^* is still homeomorphic to $B(k)$. Finally, if $k = \aleph_0$, the whole construction simplifies greatly; we take each $C_i(c_{i-1})$ to consist of just 2 points, and C^* will now be homeomorphic to the Cantor set.

4.3. COROLLARY. *Under the above hypotheses on X, X contains a homeomorph of $B(k)$.*

For apply the theorem taking $Y = X$ and $f =$ identity mapping.

5. Sequences of ordinals.

5.1. Consider the space S of all sequences of countable ordinals. For the present, we give the ordinals the *discrete* topology; thus S is just $B(\aleph_1)$, regarded as $\prod_n \Omega_n$ where Ω_n is (for each $n = 1, 2, \cdots$) the set of all ordinals $< \omega_1$ in the discrete topology; S is, of course, given the product topology. For each limit ordinal $\alpha < \omega_1$, "choose" a sequence $b(\alpha) = (\beta_1, \beta_2, \cdots)$ of ordinals $\beta_1, \beta_2, \cdots$ such that $\beta_n < \alpha$ and $\sup_n \beta_n = \alpha$. Consider the subset $E = \{b(\alpha) \mid \alpha \text{ a limit ordinal } < \omega_1\}$ of S. Our object is to confirm the feeling that there is no "nice" way of carrying out the above choices, by proving (Theorem 5 below) that E is never a Borel subset of S. The same result will then follow when we give the ordinals their usual topology (see 5.5). In what follows, all ordinals used are understood to be less than ω_1.

5.2. LEMMA 5. *E is not σ-discrete.*

We give $S = B(\aleph_1)$ the usual metric in which $\rho((x_1, x_2, \cdots), (y_1, y_2, \cdots)) = 1/n$ when $x_i = y_i$ for all $i < n$ but $x_n \neq y_n$. Suppose E is σ-discrete; then (Lemma 1, 2.1) $E = \cup \{D_n \mid n = 1, 2, \cdots\}$ where D_n is ϵ_n-discrete for some

$\epsilon_n > 0$. For each $i = 1, 2, \cdots$, put $E_i = \bigcup \{D_n \mid \epsilon_n > 1/i\}$; thus $\bigcup_i E_i = E$. We first show:

(1) Given $\beta_1, \beta_2, \cdots, \beta_i$, there are at most $\aleph_0$ points of E_i beginning with $(\beta_1, \beta_2, \cdots, \beta_i)$.

For let P be the set of those points $p \in E_i$ which have the form

$$p = (\beta_1, \beta_2, \cdots, \beta_i, p_{i+1}, \cdots).$$

Each $p \in P$ is in some $D_{n(p)}$ with $\epsilon_{n(p)} > 1/i$. Now if $p, p' \in P$ and $p \neq p'$, then $\rho(p, p') \leqq 1/(i+1)$, and therefore $n(p) \neq n(p')$, whence P is countable.

For each ordinal α and each positive integer i, write $E_i(\alpha) =$ subset of E_i consisting of those points of E_i whose first i co-ordinates are all $\leqq \alpha$. We show

(2) each $E_i(\alpha)$ is countable.

For there are at most $\aleph_0$ possibilities for $(\beta_1, \beta_2, \cdots, \beta_i)$ with $\beta_1, \cdots, \beta_i$ all $\leqq \alpha$; and, from (1), each such possibility gives rise to at most $\aleph_0$ points of E_i.

For each $p = (\beta_1, \beta_2, \cdots) \in E$ we write $a(p) = \sup_n \beta_n$; thus $a(b(\alpha)) = \alpha$ for each limit ordinal α. We define a sequence $\alpha_1 < \alpha_2 < \cdots$ of ordinals as follows. Take $\alpha_1 = \omega_0$, say. When α_r has been defined, we note that $E_i(\alpha_r)$ is countable, by (2), so we may define an ordinal γ_{ri} $(< \omega_1)$ by: $\gamma_{ri} = \sup\{a(p) \mid p \in E_i(\alpha_r)\}$; and we take α_{r+1} to be any ordinal greater than both α_r and $\sup_i \gamma_{ri}$. This defines an increasing sequence $\{\alpha_r \mid r = 1, 2, \cdots\}$. Let $\alpha^* = \sup_r \alpha_r$; this is, of course, a limit ordinal $(< \omega_1)$, so the sequence $b(\alpha^*) = (\beta_1^*, \beta_2^*, \cdots)$ belongs to E, and therefore belongs to E_n for some positive integer n. Further, we have $\max\{\beta_1^*, \cdots, \beta_n^*\} < \alpha^*$, and therefore $< \alpha_s$ for some positive integer s. This means that $b(\alpha^*) \in E_n(\alpha_s)$. By construction, therefore, $\alpha_{s+1} > \gamma_{sn} \geqq a(b(\alpha^*)) = \alpha^* > \alpha_{s+1}$, giving a contradiction.

5.3. LEMMA 6. *Every separable subset of E is countable.*

Let C be any countable subset of E; say $C = \{b(\alpha_m) \mid m = 1, 2, \cdots\}$. Put $\alpha^* = \sup_m \alpha_m$, and let F denote the set of all points of S all of whose co-ordinates are $\leqq \alpha^*$. Then F is closed in S, so $\bar{C}$ (relative to E) $\subset E \cap F$. The map $a \mid E \cap F$ is a 1-1 correspondence between $E \cap F$ and a countable set of ordinals (all $\leqq \alpha^*$); so $E \cap F$ is countable, whence $\bar{C}$ is countable also.

5.4. THEOREM 5. *E is not Borel (in S), though every separable subset of E is an absolute F_σ.*

For, if E were absolutely Borel, Theorem 2 (equivalence of (4) and (7)) would contradict Lemmas 5 and 6.

5.5. Corollary. *E is not Borel in $\Omega^{\aleph_0}$, where Ω is the space of countable ordinals in the usual order-topology.*[5]

For the space S above is $\Omega^{\aleph_0}$ with a finer topology. Hence if E were Borel in $\Omega^{\aleph_0}$, it would be Borel in S.

6. Spaces of weight $\aleph_1$.

6.1. It was shown in [4, Th. 11, 5.5] that every absolute Borel set of weight $\aleph_1$ is Borel isomorphic (and in fact generalized homeomorphic) to one of the following 4 spaces: (1) $\aleph_1 B(1)$, (2) $\aleph_1 B(1) + B(\aleph_0)$, (3) $\aleph_1 B(\aleph_0)$, (4) $B(\aleph_1)$. Here "+" denotes discrete union, $B(1)$ is a 1-point space, and the notation kS (where S is a space and k a cardinal) means the discrete union of k copies of S. It was further shown that no two of these 4 spaces are Borel isomorphic, with the possible exception of (3) and (4). We can now complete the classification of the absolute Borel sets of weight $\aleph_1$ by proving:

Theorem 6. *$\aleph_1 B(\aleph_0)$ and $B(\aleph_1)$ are not Borel isomorphic.*

Proof. Suppose, on the contrary, that f is a Borel isomorphism of $R = \aleph_1 B(\aleph_0)$ onto $S = B(\aleph_1)$. Now R is the discrete union of $\aleph_1$ copies B_λ $(\lambda < \omega_1)$ of $B(\aleph_0)$. Consider the subset E of S which was studied in the previous section; write $D = f^{-1}(E)$, $D_\lambda = D \cap B_\lambda$. Thus $f(D_\lambda) = f(B_\lambda) \cap E$.

Now B_λ is a separable absolute Borel set; hence [4, p. 20] $f(B_\lambda)$ is separable, and so $f(D_\lambda)$ is a separable subset of E. By Lemma 6 (5.3) $f(D_\lambda)$ is countable; thus so is D_λ; enumerate its points as $\{d_{\lambda n} \mid n \in N_\lambda\}$, where N_λ is a suitable set of positive integers (possibly finite or empty). For each positive integer n, put $F_n = \{d_{\lambda n} \mid \lambda < \omega_1, n \in N_\lambda\}$; then F_n is closed in R because it meets each B_λ in at most one point. Hence $D = \bigcup_n F_n$ is F_σ in R, and therefore $f(D)$ must be Borel in S. But $f(D) = E$, which is not Borel, by Theorem 5; and this gives the desired contradiction.

6.2. It would be interesting to know whether every subset of E is Borel in E. If so, E would be an example of a space which is Borel isomorphic to an absolute Borel set (in fact to a discrete set of $\aleph_1$ points) without itself being absolutely Borel.

University of Rochester,
Rochester, N. Y.

REFERENCES.

[1] C. Kuratowski, *Topologie* I (2nd ed.), Warsaw, 1948.

[2] E. Michael, "Local properties of topological spaces," *Duke Mathematical Journal*, vol. 21 (1954), pp. 163-172.

[3] A. H. Stone, "Cardinals of closed sets," *Mathematika*, vol. 6 (1959), pp. 99-107.

[4] ———, "Non-separable Borel sets," *Rozprawy Matematyczne* 28, Warsaw, 1962.

[5] ———, "Absolute F_σ-spaces," *Proceedings of the American Mathematical Society*, vol. 13 (1962), pp. 495-499.

[6] ———, "Kernel constructions and Borel sets," *Transactions of the American Mathematical Society*, vol. 107 (1963), pp. 58-70.

DUALITY AND RADON TRANSFORM FOR SYMMETRIC SPACES.*

By S. HELGASON.[1]

1. **Introduction.** The classical Radon transform for the Euclidean space $\boldsymbol{R}^n$ associates to each continuous function f of compact support the integrals of f over the hyperplanes in $\boldsymbol{R}^n$. If f is of class C^∞ then it can be recovered from these integrals by means of a simple formula. This inversion formula amounts to a decomposition of f into plane waves, that is into functions which have parallel planes as level surfaces. A plane wave is just a function of one variable so the decomposition mentioned can sometimes reduce a problem for functions in n variables to a similar problem for functions in one variable.

Now let S be a Riemannian globally symmetric space of the noncompact type and $G = I_0(S)$ the largest connected group of isometries of S in the compact open topology. The group G contains maximal unipotent subgroups (all mutually conjugate) and their orbits in S are called horocycles. Each horocycle has dimensional equal to $\dim S - \operatorname{rank} S$. The space $\hat{S}$ of all horocycles in S is another homogeneous space acted on transitively by G and will be called the dual of S in view of the analogy with the projective duality between points and hyperplanes in $\boldsymbol{R}^n$. Although totally geodesic subspaces in S are the closest analogy to planes in $\boldsymbol{R}^n$, the horocycles serve better in generalizing the Radon transform because S is a disjoint union of parallel horocycles. The "polar coordinate" decomposition of S (Prop. 3.7) has an analogue for $\hat{S}$ (Prop. 3.6). Moreover, if we consider the isotropy groups for the action of G on S and $\hat{S}$, respectively, the orbit spaces of these isotropy groups have a dual structure (Theorem 3.8). The algebras $\boldsymbol{D}(S)$ and $\boldsymbol{D}(\hat{S})$ of G-invariant differential operators on S and $\hat{S}$, respectively, also have similar properties.

By analogy with the classical Radon transform, the Radon transform of a continuous function f of compact support on S is defined to be the function $\hat{f}$ on $\hat{S}$ such that for $\xi \in \hat{S}$, $\hat{f}(\xi)$ is the integral of f over ξ. This mapping $f \to \hat{f}$ is one-to-one and induces an isomorphism $D \to \hat{D}$ of $\boldsymbol{D}(S)$ into $\boldsymbol{D}(\hat{S})$ described in Theorem 5.3. On the other hand there is a natural dual to the

* Received August 1, 1963.

[1] This work was supported in part by the National Science Foundation, NSF GP-149.

transform $f \to \hat{f}$. This dual transform associates to a continuous function ψ on $\hat{S}$ the function $\check{\psi}$ on S such that for $p \in S$, $\check{\psi}(p)$ is the average of ψ over the (compact) set of horocycles passing through p. A natural problem is now to relate f and $(\hat{f})^{\vee}$. This is solved by Theorem 6.1 for the case when G is a complex Lie group, the principal tool in the proof being the Plancherel formula for the group G. This theorem can be interpreted as a decomposition of a function on S into plane waves.

In § 7 the preceding results (G assumed complex) are applied to invariant differential equations on S. It is shown that a differential equation $Du = f$ where $D \in \boldsymbol{D}(S)$ and f is a given function can be reduced to a differential equation with constant coefficients on a Euclidean space (and is thus, in principle, solvable). The Euclidean space in question is a maximal flat totally geodesic submanifold of S. In [13b] a different method is given, applying also to real G.

A short summary of this paper is published in [13a].

2. Notation. As usual, $\boldsymbol{R}$ and $\boldsymbol{C}$ denote the fields of real and complex numbers respectively. The identity element of a group will be denoted by e. If A is a group and B a subgroup the system of left cosets aB will be denoted by A/B. The subgroup B will, when viewed as an element in A/B, be denoted $\{B\}$. For each $x \in A$, the mapping $aB \to xaB$ of A/B onto itself will be denoted $\tau(x)$; the left translation $a \to xa$ and the right translation $a \to ax$ will be denoted L_x (or $L(x)$) and R_x (or $R(x)$) respectively. If M is a topological space, $C(M)$ (respectively, $C_c(M)$) denotes the space of continuous functions (respectively, continuous functions of compact support) on M. If $L(M)$ is any linear space of functions f on M and D is an endomorphism of $L(M)$ then $[Df](p)$ and sometimes $D_p(f(p))$ denotes the value of the function Df at p. If M_1 and M_2 are two topological spaces and τ a homeomorphism of M_1 onto M_2 we put f^τ for the composite function $f \circ \tau^{-1}$. If $M_1 = M_2 = M$ and the maps $f \to f^\tau$ and $f \to f^{\tau^{-1}}$ leave $L(M)$ invariant we define D^τ by $D^\tau f = (Df^{\tau^{-1}})^\tau$ for $f \in L(M)$. If $D^\tau = D$ then D is called invariant under τ. If M is a manifold, $C^\infty(M)$ (respectively $C^\infty{}_c(M)$) denotes the space of differentiable functions (respectively, differentiable functions with compact support) on M. If D is a differential operator on M and τ a diffeomorphism of M then D^τ is also a differential operator on M.

Lie group will be denoted by capital Roman letters and their Lie algebras by corresponding lower case German letters. The adjoint representation of a Lie group G (respectively, Lie algebra $\mathfrak{g}$) will be denoted Ad_G (respectively, $\mathrm{ad}_{\mathfrak{g}}$). The subscripts are ommited when no confusion is likely.

3. The dual space of a symmetric space. Let S be a symmetric space (that is a Riemannian globally symmetric space), and let $I_0(S)$ denote the largest connected group of isometries of S in the compact open topology. It will always be assumed that S is of *the noncompact type,* that is $I_0(S)$ is semisimple and has no compact normal subgroup $\neq \{e\}$. The group $I_0(S)$ is generated by all products $s_p \circ s_q$ where s_p and s_q denote the geodesic symmetries with respect to the points $p, q \in S$. If we change the Riemannian structure on S to another one which is also invariant under $I_0(S)$ the geodesics remain the same, the geodesic symmetries s_p are still isometries so $I_0(S)$ remains the same. In view of this fact it is natural to choose the Riemannian structure as follows: For each $p \in S$ let K_p denote the isotropy subgroup of $G = I_0(S)$ at p, let $\mathfrak{k}_p$ and $\mathfrak{g}_0$ denote the corresponding Lie algebras and let $\mathfrak{p}_p$ denote the orthogonal complement of $\mathfrak{k}_p$ in $\mathfrak{g}_0$ with respect to the Killing form B of $\mathfrak{g}_0$. Then B is strictly positive definite on $\mathfrak{p}_p$ and the differential of the mapping $g \rightarrow g \cdot p$ of G onto S induces a linear one-to-one mapping of $\mathfrak{p}_p$ onto the tangent space S_p. The Riemannian structure on S is chosen such that this mapping of $\mathfrak{p}_p$ onto S_p is an isometry for each $p \in S$.

Let l denote the rank of S. Then S contains flat, totally geodesic submanifolds of dimension l. Let us simply call these manifolds *planes* in S. Now fix a point $o \in S$, let π denote the mapping $g \rightarrow g \cdot o$ of G onto S and let θ denote the automorphism $g \rightarrow s_o g s_o$ of G as well as the corresponding automorphism of $\mathfrak{g}_0$. The Exponential mapping Exp at o is a diffeomorphism of S_o onto S. Let E be any plane containing o and let $\mathfrak{h}_{\mathfrak{p}_0}$ denote the (maximal abelian) subspace of vectors $X \in \mathfrak{p}_0$ for which $\mathrm{Exp}(d\pi X) \in E$. The subalgebra $\mathfrak{h}_{\mathfrak{p}_0} \subset \mathfrak{g}_0$ is often called a *Cartan subalgebra* for the space S. Let K stand for the isotropy group K_o of o and let $A_\mathfrak{p}$ denote the analytic subgroup of G corresponding to $\mathfrak{h}_{\mathfrak{p}_0}$. An element $H \in \mathfrak{h}_{\mathfrak{p}_0}$ is called regular if its centralizer in $\mathfrak{p}_0$ equals $\mathfrak{h}_{\mathfrak{p}_0}$. Let $\mathfrak{h}'_{\mathfrak{p}_0}$ denote the set of regular elements in $\mathfrak{h}_{\mathfrak{p}_0}$. The connected components of $\mathfrak{h}'_{\mathfrak{p}_0}$ are called the Weyl chambers in $\mathfrak{h}_{\mathfrak{p}_0}$. Select any Weyl chamber C in $\mathfrak{h}_{\mathfrak{p}_0}$. Then the dual space of $\mathfrak{h}_{\mathfrak{p}_0}$ can be ordered by calling a linear function λ on $\mathfrak{h}_{\mathfrak{p}_0}$ *positive* if $\lambda(H) > 0$ for all $H \in C$. This ordering gives rise to an Iwasawa decomposition of G, $G = KA_\mathfrak{p}N$ where N is a connected nilpotent subgroup of G (see e.g. [13], Ch. VI). The Lie algebra $\mathfrak{n}_0$ of N is given by $\mathfrak{n}_0 = \sum_{\lambda > 0} \mathfrak{g}_{0\lambda}$, where for each linear function λ on $\mathfrak{h}_{\mathfrak{p}_0}$ we put $\mathfrak{g}_{0\lambda} = \{X \in \mathfrak{g}_0 \mid [H, X] = \lambda(H)X,\ H \in \mathfrak{h}_{\mathfrak{p}_0}\}$.

The group N depends on the choices of o, E and C. However, since all pairs (o, E) where $o \in E$ are conjugate under G and all Weyl chambers C in $\mathfrak{h}_{\mathfrak{p}_0}$ are conjugate under K it follows that the nilpotent subgroups of the various

Iwasawa decompositions of G are all conjugate in G, i.e. all of the form gNg^{-1} $(g \in G)$.

Definition. A *horocycle* in S is an orbit of a subgroup of the form gNg^{-1}, g being any element in G.

Remark. Identifying G with its adjoint group, the groups gNg^{-1} are the maximal unipotent subgroups of G ([1], [7], [17], [21]). Thus the horocycles in S can be defined quite directly, that is without invoking the Iwasawa decomposition.

We shall now turn the set of horocycles into a manifold and study this manifold in some detail. First we introduce some notation.

Let M and M', respectively, denote the centralizer and normalizer of $\mathfrak{h}_{\mathfrak{p}_0}$ in K. Let W denote the (finite) Weyl group M'/M. Let $\mathfrak{h}_0$ be any maximal abelian subalgebra of $\mathfrak{g}_0$ containing $\mathfrak{h}_{\mathfrak{p}_0}$, let $\mathfrak{g}$ denote the complexification of $\mathfrak{g}_0$ and $\mathfrak{h}$ the subspace of $\mathfrak{g}$ spanned by $\mathfrak{h}_0$. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Let Δ denote the set of nonzero roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and let $\mathfrak{h}_{\mathfrak{k}_0} = \mathfrak{h} \cap \mathfrak{k}_0$, $\mathfrak{h}^* = \mathfrak{h}_{\mathfrak{p}_0} + i\mathfrak{h}_{\mathfrak{k}_0}$. All the roots in Δ are real on $\mathfrak{h}^*$. Let C^* be any Weyl chamber in $\mathfrak{h}^*$ whose closure contains the Weyl chamber C in $\mathfrak{h}_{\mathfrak{p}_0}$. We order the dual space of $\mathfrak{h}^*$ by means of the Weyl chamber C^*. Then if $\bar{\alpha}$ denotes the restriction to $\mathfrak{h}_{\mathfrak{p}_0}$ of a root $\alpha \in \Delta$ then $\alpha > 0 \Rightarrow \bar{\alpha} \geqq 0$ and $\bar{\alpha} > 0 \Rightarrow \alpha > 0$. The set Δ^+ of positive roots in Δ is a disjoint union, $\Delta^+ = P_+ \cup P_-$, where $\alpha \in \Delta^+$ belongs to P_+ or P_- respectively according to whether $\bar{\alpha} > 0$ or $\bar{\alpha} = 0$. Let $\rho = \frac{1}{2} \sum_{\alpha \in P_+} \alpha$. Now $H \in \mathfrak{h}'_{\mathfrak{p}_0}$ if and only if $\alpha(H) \neq 0$ for all $\alpha \in P_+$. The set $\{H \in \mathfrak{h}_{\mathfrak{p}_0} \mid \alpha(H) > 0 \text{ for } \alpha \in P_+\}$ is therefore a subset of $\mathfrak{h}'_{\mathfrak{p}_0}$ which is convex, hence connected, and contains C. Consequently

$$C = \{H \in \mathfrak{h}_{\mathfrak{p}_0} \mid \alpha(H) > 0 \text{ for } \alpha \in P_+\}.$$

LEMMA 3.1. *Let H be any element in C. Then*

$$N = \{z \in G \mid \lim_{t \to \infty} \exp(-tH) z \exp(tH) = e\}. \tag{1}$$

This lemma is presumably known (cf. [16]) but no proof seems available. Denoting the right hand side of (1) by N_H the previous expression for $\mathfrak{n}_0$ shows that $N \subset N_H$. The converse requires some matrix computation. As is well known (see e.g. [13], p. 223) there exists a basis of $\mathfrak{g}$ such that the matrices representing $\mathrm{ad}_{\mathfrak{g}}$ have the following properties:

$\mathrm{ad}(\mathfrak{k}_0)$: skew Hermitian;

$\mathrm{ad}(\mathfrak{n}_0)$: lower triangular with zeros in the diagonal;

$\operatorname{ad}(\theta\mathfrak{n}_0)$: upper triangular with zeros in the diagonal;

$\operatorname{ad}_{\mathfrak{g}}(H)$ $(H\in\mathfrak{h})$ is a diagonal matrix with diagonal $\lambda_1(H),\lambda_2(H),\cdots$, where $\lambda_1,\lambda_2,\cdots$ are all the roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ in increasing order (here the root 0 is counted with multiplicity $\dim\mathfrak{h}$).

Now suppose $ka\in N_H$ for some $k\in K$, $a\in A_{\mathfrak{p}}$. Then

$$(2)\qquad \operatorname{Ad}(\exp(-tH)k\exp(tH))\to\operatorname{Ad}(a^{-1})$$

for $t\to\infty$. Consider the matrices in (2),

$$(3)\qquad (\operatorname{Ad}(\exp(-tH)k\exp(tH)))_{ij}=\exp((\lambda_j-\lambda_i)(tH))k_{ij}$$

where $(\operatorname{Ad}(k))_{ij}=k_{ij}$. Then it follows from (2) that $k_{ij}=0$ for $i<j$ and $k_{ii}>0$ for all i. But being also unitary, $\operatorname{Ad}(k)=e$, so we find $ka=e$; q. e. d.

LEMMA 3.2. $(N\theta(N))\cap K=\{e\}$.

Proof. Suppose $n_1\theta(n_2)\in K$ for some $n_1,n_2\in N$. Let $\operatorname{Ad}_G(n_1\theta(n_2))_{ij}=u_{ij}$ in terms of the basis of $\mathfrak{g}$ used above. Clearly $u_{11}=1$ and therefore $u_{1j}=0$ for $j\neq 1$, the matrix (u_{ij}) being unitary. This implies $\operatorname{Ad}_G(\theta(n_2))_{1j}=0$ for $j>1$. Hence $u_{22}=1$ and $u_{2j}=0$ for $j\neq 2$. By induction we find that $u_{ij}=\delta_{ij}$ so $n_1\theta(n_2)$ lies in the center of G. But this center lies in K and G acts effectively so the lemma is reduced to proving $N\cap\theta(N)=\{e\}$. This however is clear from the matrix representation already used.

LEMMA 3.3. *The plane E is orthogonal to the horocycle $N\cdot o$ at o.*

Let $\mathfrak{q}_0$ denote the subspace of $\mathfrak{p}_0$ which $d\pi$ maps onto the tangent space to the orbit $N\cdot o$ at o. This tangent space consists of the vectors $d\pi(X)$ as X runs through $\mathfrak{n}_0$. Let $X_1\in\mathfrak{q}_0$, $H\in\mathfrak{h}_{\mathfrak{p}_0}$; then $d\pi(X_1)=d\pi(X)$ for some $X\in\mathfrak{n}_0$. Hence $X_1-X\in\mathfrak{k}_0$ so $B(X_1,H)=B(X,H)=0$, proving the lemma.

THEOREM 3.4. *The group G acts transitively on the set of horocycles in S. The subgroup of G which maps the horocycle $N\cdot o$ into itself equals MN.*

Proof. Let $g,h\in G$ and consider the orbits $gNg^{-1}h\cdot o$ and $N\cdot o$. Writing $h^{-1}g=kan$ and using $aNa^{-1}=N$ we find that $hkN\cdot o=gNg^{-1}h\cdot o$ so G is transitive.

Now $mNm^{-1}\subset N$ for each $m\in M$ so MN is a group which maps $N\cdot o$ into itself. On the other hand let $g\in G$ be such that $gN\cdot o=N\cdot o$. Writing $g=kan$ we conclude that $kann_1=k_1$ for some $n_1\in N$, $k_1\in K$. Hence $a=e$ so $kN\cdot o=N\cdot o$. Lemma 3.3 now implies that $\operatorname{Ad}_G(k)\mathfrak{h}_{\mathfrak{p}_0}\subset\mathfrak{h}_{\mathfrak{p}_0}$, that is, $k\in M'$. Suppose $k\notin M$. Then the Weyl group element $s=kM$ does not leave C invariant so given $H_s\in s^{-1}C$ there exists a root $\alpha\in P_+$ such that $\alpha(H_s)<0$.

Then the hyperplane $\alpha(H)=0$ in $\mathfrak{h}_{\mathfrak{p}_0}$ separates the Weyl chambers C and $s^{-1}C$. Select $X\neq 0$ in $\mathfrak{n}_0$ such that $[H,X]=\alpha(H)X$ for $H\in\mathfrak{h}_{\mathfrak{p}_0}$. Since $\alpha(s^{-1}H)<0$ for $H\in C$ we have $\mathrm{Ad}(k)X\in\theta(\mathfrak{n}_0)$ so $kNk^{-1}\cap\theta(N)\neq\{e\}$. But $kN\cdot o=N\cdot o$ implies $kNk^{-1}\subset NK$ so Lemma 3.2 gives a contradiction. Thus $k\in M$ and Theorem 3.4 is proved.

Since M is compact and N is closed in G, the subgroup MN is closed in G.

Definition. The set of horocycles in S, with the differentiable structure of G/MN, will be called the *dual space of* S and denoted by $\hat{S}$.

Although the differentiable structure of $\hat{S}$ involves a choice of o, E and C it is clear that a different choice only changes the subgroup MN into a conjugate subgroup $g(MN)g^{-1}$ and this gives the same differentiable structure on $\hat{S}$.

Let $t\to\gamma(t)$ $(t\in\boldsymbol{R})$ be any geodesic in S. Let s_τ denote the geodesic symmetry of S with respect to the point $\gamma(\tau)$ and put $T_t=s_{t/2}s_0$. Then $T_t(t\in\boldsymbol{R})$ is a one-parameter subgroup of G. Its element are called *transvections* along the geodesic γ. Two horocycles ξ_1 and ξ_2 are called parallel if there exists a geodesic γ intersecting ξ_1 and ξ_2 under a right angle and such that $T(\xi_1)=\xi_2$ for a suitable transvection T along γ. If $\xi\in\hat{S}$, $\{\xi\}$ shall denote the family of horocycles parallel to ξ.

LEMMA 3.5. *For a fixed $g\in G$ the orbits of the group gNg^{-1} form a parallel family of horocycles. Each parallel family of horocycles arises in this fashion.*

Proof. Since $G=NA_{\mathfrak{p}}K$ each orbit of N intersects E. If $a\in A_{\mathfrak{p}}$, the orbit $N\cdot(a\cdot o)$ equals $aN\cdot o$ which is parallel to $N\cdot o$ (Lemma 3.3). Since G acts transitively on the set $\{\{\xi\}\mid\xi\in\hat{S}\}$ the lemma follows.

Consider now the set of horocycles which pass through o. In view of Theorem 3.4 this set is in a natural correspondence with the coset space K/M, which again is in an obvious natural correspondence with the set of all Weyl chambers in all maximal abelian subspaces of $\mathfrak{p}_0$.

PROPOSITION 3.6. *The mapping $\psi:(kM,a)\to ka(N\cdot o)$ is a diffeomorphism of $K/M\times A_{\mathfrak{p}}$ onto $\hat{S}$.*

Proof. Let $k\in K$ and put $\xi=k(N\cdot o)$; since $(kak^{-1})kNk^{-1}\cdot o$ is parallel to ξ it is clear that the parallel family $\{\xi\}$ can be written $\{kaN\cdot o\mid a\in A_{\mathfrak{p}}\}$. Thus ψ is onto. Next suppose $k_1,k_2\in K$, $a_1,a_2\in A_{\mathfrak{p}}$ such that $k_1a_1N\cdot o=k_2a_2N\cdot o$. By Theorem 3.4, $(k_2a_2)^{-1}k_1a_1\in MN$ and since $aNa^{-1}=N$, $aMa^{-1}=M$ $(a\in A_{\mathfrak{p}})$ we obtain $k_2^{-1}k_1a_1a_2^{-1}\in MN$ so $k_1M=k_2M$ and $a_1=a_2$.

Thus ψ is one-to-one. Now let $\mathfrak{m}_0$ denote the Lie algebra of M and let $\mathfrak{l}_0$ denote the orthogonal complement of $\mathfrak{m}_0$ in $\mathfrak{k}_0$ with respect to B. Consider the natural projections

$$\pi_0\colon K \to K/M, \quad \pi\colon G \to G/K, \quad \hat{\pi}\colon G \to G/MN$$

and let $\tau(x)$ be as described in Section 2. The differentials $d\pi_0$ and $d\hat{\pi}$ map $\mathfrak{l}_0$ and $\mathfrak{l}_0 + \mathfrak{h}_{\mathfrak{p}_0}$ respectively, isomorphically onto the tangent spaces $(K/M)_{\pi_0(e)}$ and $(G/MN)_{\pi^\wedge(e)}$. Let $a \in A_\mathfrak{p}$, $k \in K$, $H \in \mathfrak{h}_{\mathfrak{p}_0}$, $L \in \mathfrak{l}_0$. Then for $t \in \boldsymbol{R}$,

$$\psi(k(\exp tL)M, a) = ka\exp(\mathrm{Ad}(a^{-1})tL)MN,$$
$$\psi(kM, a\exp tH) = ka(\exp tH)MN,$$

so

$$d\psi(d\tau(k)\,d\pi_0(L), d\tau(a)H) = d\tau(ka)\,(d\hat{\pi}(\mathrm{Ad}(a^{-1})L + H)).$$

The right hand side is $\neq 0$ unless $\mathrm{Ad}(a^{-1})L + H \in \mathfrak{m}_0 + \mathfrak{n}_0$. But this happens only if $L = H = 0$ so ψ is regular and the proposition is proved.

In comparison we recall the following well known fact.

PROPOSITION 3.7. *The mapping* $\phi\colon (kM, a) \to kaK$ *is a differentiable mapping of* $(K/M) \times A_\mathfrak{p}$ *onto* S *and a regular w-to-one mapping of* $K/M \times A'_\mathfrak{p}$ *onto* S'.

Here w denotes the order of the Weyl group M'/M, $A'_\mathfrak{p} = \exp \mathfrak{h}'_{\mathfrak{p}_0}$ and S' is the set of points in S which lie on only one plane through o.

We shall now describe the double coset spaces $K\backslash G/K$ and $MN\backslash G/MN$ which of course can be regarded as the orbit spaces for the action of K on G/K and of MN on G/MN. Let $A_\mathfrak{p}/W$ be the space obtained from $A_\mathfrak{p}$ by identifying points which are conjugate under W.

THEOREM 3.8. *The following relations are natural identifications:*

(i) $K\backslash G/K = A_\mathfrak{p}/W$;

(ii) $MN\backslash G/MN = A_\mathfrak{p} \times W$.

Proof. The first relation is well known and is proved as follows: Each $g \in G$ can be written $g = k_1ak_2$ $(k_1, k_2 \in K, a \in A_\mathfrak{p})$. If $g = k'_1a'k'_2$ is another such representation one finds by applying the automorphism θ, $ka^2k^{-1} = (a')^2$, where $k \in K$. As is known (see e.g. [13], p. 245) this implies that a^2 and $(a')^2$ (and therefore a and a') are conjugate under the Weyl group. This proves (i).

In order to prove (ii) we recall a result of Bruhat-Harish-Chandra [9] which identifies the double coset space $MNA_\mathfrak{p}\backslash G/MNA_\mathfrak{p}$ with the Weyl group.

More precisely, each $g \in G$ can be written

$$(4) \qquad g = m_1 n_1 a_1 m_w a_2 n_2 m_2 \quad (m_1, m_2 \in M, n_1, n_2 \in N, a_1, a_2 \in A_{\mathfrak{p}})$$

where $m_w \in M'$ is uniquely determined $(\operatorname{mod} M)$. Since $M'A_{\mathfrak{p}}(M')^{-1} \subset A_{\mathfrak{p}}$, (4) can be written

$$(5) \qquad g = m_1 n_1 m_w a n_2 m_2$$

and (ii) therefore amounts to, that not only is $m_w M$ uniquely determined by g, but that the same is true of a. For this we need a lemma.

LEMMA 3.9. *Let C and C' be two Weyl chambers in $\mathfrak{h}_{\mathfrak{p}_0}$ and $G = KA_{\mathfrak{p}}N$, $G = KA_{\mathfrak{p}}N'$ the corresponding Iwasawa decompositions. Then*

$$(NN') \cap (MA_{\mathfrak{p}}) = \{e\}.$$

Proof. We first consider the case $C' = -C$; here $N' = \theta(N)$. Let n_1, $n_2 \in N$, $m \in M$, $a \in A_{\mathfrak{p}}$ such that

$$(6) \qquad n_1\theta(n_2) = ma.$$

Now we apply θ and use $ma = am$. Then

$$\theta(n_1) n_2 n_1 \theta(n_2) = m^2.$$

But $m\theta(N)m^{-1} = \theta(N)$ so we get from Lemma 3.2, $m^2 = e$ and $n_2 n_1 = e$. Let b denote the unique square root of a in $A_{\mathfrak{p}}$. Since m commutes with b we obtain from (6), $b^{-1}n_1\theta(n_1^{-1})b^{-1} = m$, or $s\theta(s^{-1}) = m$ where $s = b^{-1}n_1$. Writing $s = (\exp X)k$ where $X \in \mathfrak{p}_0$ and $k \in K$ it follows that $m = \exp 2X$ so $m^2 = e$ implies $m = e$. Now, in terms of the basis of $\mathfrak{g}$ used earlier, $\operatorname{Ad}_G(n_1)$ is lower triangular and $\operatorname{Ad}_G(\theta(n_1^{-1}))$ is upper triangular. Their diagonals are 1 and their product is (by (6)) a diagonal matrix. It follows that $n_1 = e$, so the relation $(N\theta(N)) \cap (MA_{\mathfrak{p}}) = \{e\}$ is proved.

The proof of the lemma will be finished by proving $N' \subset N\theta(N)$. For each $\alpha \in \Delta$, let $\mathfrak{g}^\alpha$ denote the root subspace

$$\mathfrak{g}^\alpha = \{X \in \mathfrak{g} \mid [H, X] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}.$$

Let P' denote the set of roots in Δ which are > 0 on C'. Then if we put $\mathfrak{n}' = \sum_{\alpha \in P'} \mathfrak{g}^\alpha$, the Lie algebra of N' is $\mathfrak{n}'_0 = \mathfrak{n}' \cap \mathfrak{g}_0$. If $-P_+$ denotes the set of negatives of the roots in P_+ it is clear that $P' \subset P_+ \cup (-P_+)$ so

$$\mathfrak{n}' = \sum_{\alpha \in P' \cap P_+} \mathfrak{g}^\alpha + \sum_{\alpha \in P' \cap (-P_+)} \mathfrak{g}^\alpha.$$

Taking intersections with $\mathfrak{g}_0$ we obtain the direct decomposition

$$\mathfrak{n}'_0 = \mathfrak{n}_1 + \mathfrak{n}_2.$$

Here $\mathfrak{n}_1 = \mathfrak{n}'_0 \cap \mathfrak{n}_0$ so it is a subalgebra of $\mathfrak{n}'_0$. Now if $\beta_1 < \beta_2 < \cdots < \beta_p$ are the roots in P' in increasing order, consider the space

$$\mathfrak{n}^{(k)} = \left(\sum_{k \leqq i \leqq p} \mathfrak{g}^{\beta_i}\right) \cap \mathfrak{g}_0 \qquad (1 \leqq k \leqq p).$$

Then $\mathfrak{n}^{(k)}$ is an ideal in $\mathfrak{n}'_0$, $[\mathfrak{n}'_0, \mathfrak{n}^{(k)}] \subset \mathfrak{n}^{(k+1)}$, $\mathfrak{n}^{(1)} = \mathfrak{n}'_0$, and $\mathfrak{n}^{(k)}$ is the direct sum of its intersections with $\mathfrak{n}_1$ and $\mathfrak{n}_2$. Using a result of [10], p. 736 we conclude that each $n \in N'$ can be written $n = \exp X \exp Y$ where $X \in \mathfrak{n}_1$, $Y \in \mathfrak{n}_2$. Hence $n \in N\theta(N)$ and Lemma 3.9 is proved.

Passing now to the proof of the theorem consider a second decomposition $g = m_3 n_3 m_w a' n_4 m_4$ of the element g in (5). Now $m_w^{-1} N m_w = N'$ for a suitable Weyl chamber C'. The lemma shows that $a = a'$ and the theorem is proved.

4. Invariant differential operators on the space of horocycles. Let A be a Lie group, B a closed subgroup. Then $\boldsymbol{D}(A/B)$ shall denote the algebra of differential operators on A/B which are invariant under the action of A. We write $\boldsymbol{D}(A)$ instead of $\boldsymbol{D}(A/\{e\})$ and sometimes we shall also write $\boldsymbol{D}(S)$ for $\boldsymbol{D}(G/K)$ and $\boldsymbol{D}(\mathcal{S})$ for $\boldsymbol{D}(G/MN)$. As will be explained in more detail in the next section, the algebra $\boldsymbol{D}(S)$ is isomorphic to the algebra $I(\mathfrak{h}_{\mathfrak{p}_0})$ of all polynomials on $\mathfrak{h}_{\mathfrak{p}_0}$ which are invariant under the Weyl group. At present our problem is to analyze $\boldsymbol{D}(\mathcal{S})$. For this purpose it is convenient to consider $\mathcal{S}$ as a fibre bundle with base K/M, fibre $A_{\mathfrak{p}}$, the projection $p: \mathcal{S} \to K/M$ being the mapping which to each horocycle associates the parallel horocycle through o. The fibre F_k over the base point kM is the set $\{kaN \cdot o \mid a \in A_{\mathfrak{p}}\}$.

Now let U be a left invariant differential operator on the Lie group $A_{\mathfrak{p}}$. We carry this operator over on F_k by means of the diffeomorphism $a \to ka(N \cdot o)$ $(a \in A_{\mathfrak{p}})$ and obtain a differential operator U_{F_k} on F_k. Now if $f \in C(\mathcal{S})$ let $f \mid F_k$ denote the restriction of f to the fibre F_k. Consider the endomorphism D_U of $C^\infty(\mathcal{S})$ given by

$$(1) \qquad (D_U f) \mid F = U_F(f \mid F),$$

where F stands for an arbitrary fibre F_k. Now for each $g \in G$ the diffeomorphism $\tau(g)$ of $\mathcal{S}$ is fibre-preserving. In fact it is geometrically clear that each isometry g of S preserves parallelism of horocycles; analytically,

$$\tau(g)(kaN \cdot o) = k_1 a_1 a N \cdot o$$

if $gk = k_1 n_1 a_1$ so $\tau(g)$ maps the fibre $\{kaN \cdot o \mid a \in A_{\mathfrak{p}}\}$ onto the fibre $\{k_1 aN \cdot o \mid a \in A_{\mathfrak{p}}\}$. Let us now prove

$$(2) \qquad (U_F(\phi\circ\tau(g)))\circ\tau(g^{-1}) = U_{\tau(g)\cdot F}(\phi) \qquad (g\in G)$$

for each fibre F and each $\phi\in C^\infty(\tau(g)\cdot F)$. In view of the definition of the U_F, is suffices to prove (2) in the case when $F=\{aN\cdot o \mid a\in A_{\mathfrak{p}}\}$ and g maps F into itself. But such a g is necessarily of the form $g=a_1mn$ where $a_1\in A_{\mathfrak{p}}$ $m\in M$, $n\in N$. In this case, however, (2) reduces to the left invariance of U.

The operator D_U defined by (1) is invariant under each $g\in G$; in fact, using (2) we have

$$\begin{aligned}(D_U{}^{\tau(g)}f) \mid F &= (D_U f^{\tau(g^{-1})})^{\tau(g)} \mid F = ((D_U f^{\tau(g^{-1})}) \mid \tau(g^{-1})\cdot F)^{\tau(g)} \\ &= (U_{\tau(g^{-1})F}(f^{\tau(g^{-1})} \mid \tau(g^{-1})F))^{\tau(g)} = U_F(f\mid F) = (D_U f)\mid F.\end{aligned}$$

Similarly, one proves that the mapping $U\to D_U$ $(U\in \boldsymbol{D}(A_{\mathfrak{p}}))$ is a homomorphism. It is clear that D_U is a continuous linear mapping of $C^\infty{}_c(\mathcal{S})$ into itself (taken with the usual Schwartz topology) and that for each $f\in C^\infty{}_c(\mathcal{S})$ the support of D_Uf is contained in the support of f. Hence, by well-known facts (see [2] e. g. [12], p. 242) D_U is a differential operator. It is clear since $A_{\mathfrak{p}}$ is abelian that $\boldsymbol{D}(A_{\mathfrak{p}})$ is canonically isomorphic to the symmetric algebra $S(\mathfrak{h}_{\mathfrak{p}_0})$ over $\mathfrak{h}_{\mathfrak{p}_0}$.

THEOREM 4.1. *The mapping $U\to D_U$ defined by* (1) *is an isomorphism of $S(\mathfrak{h}_{\mathfrak{p}_0})$ onto $\boldsymbol{D}(\mathcal{S})$. In particular $\boldsymbol{D}(\mathcal{S})$ is commutative.*

For simplicity, put $H=MN$ and let o denote the point $\{H\}$ in G/H. If G/H were a reductive coset space we know from [12], p. 269 that $\boldsymbol{D}(G/H)$ is determined by the invariants of the linear group $d\tau(h)$ $(h\in H)$ acting on the symmetric algebra over the tangent space $(G/H)_o$. Although G/H is not in general reductive we shall nevertheless determine $\boldsymbol{D}(G/H)$ by means of the mentioned invariants.

Consider the subspaces $\mathfrak{k}_0$, $\mathfrak{p}_0$, $\mathfrak{n}_0$, $\mathfrak{h}_{\mathfrak{p}_0}$, $\mathfrak{h}_{\mathfrak{k}_0}$, $\mathfrak{m}_0$, $\mathfrak{l}_0$ introduced already and their respective complexifications, $\mathfrak{k}$, $\mathfrak{p}$, $\mathfrak{n}$, $\mathfrak{h}_{\mathfrak{p}}$, $\mathfrak{h}_{\mathfrak{k}}$, $\mathfrak{m}$, $\mathfrak{l}$ in $\mathfrak{g}$. For each root subspace $\mathfrak{g}^\alpha$ $(\alpha\in\Delta)$ select a nonzero vector $X_\alpha\in\mathfrak{g}^\alpha$. Then

$$(3) \qquad \mathfrak{l}=\sum_{\alpha\in P_+}\boldsymbol{C}(X_\alpha+\theta X_\alpha), \quad \mathfrak{m}=\mathfrak{h}_{\mathfrak{k}}+\sum_{\alpha\in P_-}(\mathfrak{g}^\alpha+\mathfrak{g}^{-\alpha}), \quad \mathfrak{n}=\sum_{\alpha\in P_+}\mathfrak{g}^\alpha.$$

Let σ_0 denote the projection of $\mathfrak{g}_0$ onto $\mathfrak{l}_0+\mathfrak{h}_{\mathfrak{p}_0}$ given by the direct decomposition $\mathfrak{g}_0=(\mathfrak{l}_0+\mathfrak{h}_{\mathfrak{p}_0})+(\mathfrak{m}_0+\mathfrak{n}_0)$. Let σ denote the analogous projection of $\mathfrak{g}$ onto $\mathfrak{l}+\mathfrak{h}_{\mathfrak{p}}$. Under the isomorphism

$$d\hat{\pi}\colon \mathfrak{l}_0+\mathfrak{h}_{\mathfrak{p}_0}\to (G/H)_o$$

[2] Note that in Prop. 1 in [12] p. 242, the word "finite" should be replaced by "locally finite."

the endomorphism $d\tau(h)$ $(h \in H)$ of $(G/H)_0$ corresponds to the endomorphism $\sigma_0 \circ \mathrm{Ad}_G(h)$ of $\mathfrak{l}_0 + \mathfrak{h}_{\mathfrak{p}_0}$ (cf. [13], p. 367-368). In order to determine the invariants of the group $d\tau(h)$ $(h \in H)$ acting on $(G/H)_0$ it suffices to determine the invariants of $\sigma_0 \circ \mathrm{Ad}_G(H)$ on $\mathfrak{l}_0 + \mathfrak{h}_{\mathfrak{p}_0}$ or equivalently, the invariants of $\sigma \circ \mathrm{Ad}_G(H)$ on $\mathfrak{l} + \mathfrak{h}_{\mathfrak{p}}$.

Let $\alpha_1 > \alpha_2 > \cdots > \alpha_p$ be the roots in P_+ in decreasing order and put $E_\alpha = X_\alpha + \theta X_\alpha$ $(\alpha \in P_+)$. If $H_1, \cdots, H_l$ is a basis of $\mathfrak{h}_{\mathfrak{p}_0}$ the elements in the symmetric algebra $S(\mathfrak{l} + \mathfrak{h}_{\mathfrak{p}})$ over $\mathfrak{l} + \mathfrak{h}_{\mathfrak{p}}$ can be regarded as polynomials in $H_1, \cdots, H_l$, $E_{\alpha_1}, \cdots, E_{\alpha_p}$. For each $X \in \mathfrak{m} + \mathfrak{n}$ let $d(X)$ denote the derivation of $S(\mathfrak{l} + \mathfrak{h}_{\mathfrak{p}})$ which extends the endomorphism $\sigma \circ \mathrm{ad}_{\mathfrak{g}}(X)$ of $\mathfrak{l} + \mathfrak{h}_{\mathfrak{p}}$. If $\alpha \in \Delta$, let $H_\alpha \in \mathfrak{h}$ be determined by $B(H, H_\alpha) = \alpha(H)$ $(H \in \mathfrak{h})$ and let α^θ denote the root $H \to \alpha(\theta H)$. Also put $H'_\alpha = \sigma(H_\alpha)$; if $\alpha \in P_+$, then $H'_\alpha \neq 0$ because $\mathfrak{h} = \mathfrak{h}_{\mathfrak{p}} + \mathfrak{h}_{\mathfrak{k}}$. The mapping $\alpha \to -\alpha^\theta$ is a permutation of P_+. Now, let $\alpha, \beta \in P_+$. Then

$$d(X_\beta)(E_\alpha) = \sigma([X_\beta, X_\alpha + \theta X_\alpha]) = \sigma([X_\beta, \theta X_\alpha]) = C'_\alpha\, \sigma([X_\beta, X_{\alpha^\theta}])$$

where $C'_\alpha \neq 0$. Since $[\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}] = \boldsymbol{C} H_\alpha$ we have

$$d(X_\beta)E_\alpha = \begin{cases} 0 & \text{if } \beta > -\alpha^\theta \\ C_\beta H'_\beta & \text{if } \beta = -\alpha^\theta;\ \text{here } C_\beta \neq 0. \end{cases}$$

Now suppose $q \in S(\mathfrak{l} + \mathfrak{h}_{\mathfrak{p}})$ is invariant under $\sigma \circ \mathrm{Ad}_G(H)$. Then $d(X)q = 0$ for each $X \in \mathfrak{m} + \mathfrak{n}$. We write q as a polynomial in $E_{\alpha_1}, \cdots, E_{\alpha_p}$ whose coefficients are polynomials in $H_1, \cdots, H_l$:

$$(4) \qquad q(H_1, \cdots, H_l, E_{\alpha_1}, \cdots, E_{\alpha_p}) = \sum r_{(n)}(H_1, \cdots, H_l) E_{\alpha_1}{}^{n_1} \cdots E_{\alpha_p}{}^{n_p}.$$

The derivation $d(X_{\alpha_1})$ annihilates each H_i and all $E_{\alpha_1}, \cdots, E_{\alpha_p}$ except one, say E_{α_j}, for which $-\alpha_j{}^\theta = \alpha_1$. Hence

$$0 = d(X_{\alpha_1})q = (C_{\alpha_1} n_j H'_{\alpha_1}) \sum' r_{(n)}(H_1, \cdots, H_l) E_{\alpha_1}{}^{n_1} \cdots E_{\alpha_j}{}^{n_j - 1} \cdots E_{\alpha_p}{}^{n_p},$$

where $\sum'$ ranges over those terms in (4) for which $n_j > 0$. Since $C_{\alpha_1} H'_{\alpha_1} \neq 0$ the last equation shows that $\sum'$ is an empty sum so E_{α_j} does not occur in (4). Applying $d(X_{\alpha_2})$ to (4) we see that E_{α_k} $(-\alpha_k{}^\theta = \alpha_2)$ does not occur in (4), etc. Since $\alpha \to -\alpha^\theta$ is a permutation of P_+ this proves that the right hand side of (4) lies in $S(\mathfrak{h}_{\mathfrak{p}})$.

LEMMA 4.2. *Let $I(\mathfrak{l}_0 + \mathfrak{h}_{\mathfrak{p}_0})$ denote the set of polynomials in $S(\mathfrak{l}_0 + \mathfrak{h}_{\mathfrak{p}_0})$ which are invariant under $\sigma \circ \mathrm{Ad}_G(MN)$. Then*

$$I(\mathfrak{l}_0 + \mathfrak{h}_{\mathfrak{p}_0}) = S(\mathfrak{h}_{\mathfrak{p}_0}).$$

It has been proved above that $I(\mathfrak{l}_0+\mathfrak{h}_{\mathfrak{p}_0})\subset S(\mathfrak{h}_{\mathfrak{p}_0})$. The converse inclusion follows from the fact that each $\mathrm{Ad}(m)$ $(m\in M)$ is the identity on $\mathfrak{h}_{\mathfrak{p}_0}$ and that if $n\in N$, then $\mathrm{Ad}(n)H\equiv H(\mathrm{mod}\,\mathfrak{n}_0)$ for each $H\in\mathfrak{h}_{\mathfrak{p}_0}$.

Turning now to the proof of Theorem 4.1, let $L_1,\cdots,L_p$ be a basis of $\mathfrak{l}_0$. If $a>0$ is sufficiently small, the mapping

$$(t_1,\cdots,t_{p+l})\to\exp(t_1H_1+\cdots+t_lH_l+t_{l+1}L_1+\cdots+t_{l+p}L_p)H$$

is a diffeomorphism of the cube $|t_i|<a$ onto an open neighborhood V of o in G/H. The inverse mapping is a local coordinate system on V and the coordinate vector fields $\partial/\partial t_i$ satisfy

$$(\partial/\partial t_i)_o=d\hat{\pi}(H_i),\qquad(\partial/\partial t_{j+l})_o=d\hat{\pi}(L_j),\qquad(1\leqq i\leqq l,1\leqq j\leqq p).$$

Let $D\in\boldsymbol{D}(G/H)$. Then there exists a unique polynomial P in $p+l$ variables such that

$$(5)\qquad[Df](o)=[P\left(\frac{\partial}{\partial t_1},\cdots,\frac{\partial}{\partial t_{p+l}}\right)f(\exp(t_1H_1+\cdots+t_{p+l}L_p)H)](0)$$

for all $f\in C^\infty(G/H)$. In the decomposition of P into homogeneous components let P_D denote the component of highest degree. The invariance of D of course implies

$$(6)\qquad[Df^{\tau(h)}](o)=[Df](o)\qquad(h\in H).$$

Considering the highest order derivatives in (5) one deduces fom (6) that

$$P_D(d\tau(h)d\hat{\pi}(H_1),\cdots,d\tau(h)d\hat{\pi}(L_p))=P_D(d\hat{\pi}(H_1),\cdots,d\hat{\pi}(L_p))\quad(h\in H),$$

so P_D is invariant under $d\tau(H)$. In view of Lemma 4.2, this implies that P_D is a polynomial in the basis vectors $d\hat{\pi}(H_1),\cdots,d\hat{\pi}(H_l)$ alone. Let $P(D)$ denote the corresponding polynomial in $H_1,\cdots,H_l$ and consider the differential operator $D_{P(D)}$ defined by (1). Then, if $f\in C^\infty(G/H)$,

$$(7)\quad[D_{P(D)}f](o)=[P(D)\left(\frac{\partial}{\partial t_1},\cdots,\frac{\partial}{\partial t_l}\right)f(\exp(t_1H_1+\cdots+t_lH_l)H)](0).$$

On comparison of (5) and (7) it is clear that $D-D_{P(D)}$ is of lower order than D at the point o, hence of lower order everywhere, both operators being invariant. This proves, by induction, that every $D\in\boldsymbol{D}(G/H)$ has the form $D=D_U$ for some $U\in S(\mathfrak{h}_{\mathfrak{p}_0})$. Now by (1),

$$(8)\quad\begin{aligned}[D_Uf](o)&=[U\left(\frac{\partial}{\partial t_1},\cdots,\frac{\partial}{\partial t_l}\right)f(\exp(t_1H_1+\cdots+t_lH_l)H)](0)\\&=[U\left(\frac{\partial}{\partial t_1},\cdots,\frac{\partial}{\partial t_l}\right)(f\circ\hat{\pi})(\exp(t_1H_1+\cdots+t_lH_l))](0)\end{aligned}$$

for all $U \in S(\mathfrak{h}_{\mathfrak{p}_0})$. This shows that the mapping $U \to D_U$ is one-to-one and Theorem 4.1 is proved.

5. The Radon transform. Let ξ be any horocycle in S, ds_ξ the volume element on ξ in the Riemannian structure on ξ induced by S. For $f \in C_c(S)$ we put

$$\hat{f}(\xi) = \int_\xi f(s)\, ds_\xi, \qquad \xi \in \hat{S}. \tag{1}$$

Definition. The function $\hat{f}$ on $\hat{S}$ given by (1) will be called the *Radon transform* of f.

It will be more convenient in what follows to express (1) in terms of functions on G. Let

$$\pi\colon G \to G/K, \qquad \hat{\pi}\colon G \to G/MN$$

denote the natural projections, and let

$$F = f \circ \pi, \qquad \hat{F} = \hat{f} \circ \hat{\pi} \qquad (f \in C_c(S)). \tag{2}$$

The group N preserves the measure ds_{ξ_0} on the horocycle $\xi_0 = N \cdot o$; hence we can select a left invariant measure dn on N such that the mapping $n \to n \cdot o$ of N onto ξ_0 is measure-preserving. Then

$$\hat{f}(g \cdot \xi_0) = \int_{g \cdot \xi_0} f(s)\, ds_{g \cdot \xi_0} = \int_{\xi_0} f(g \cdot s_o)\, ds_o = \int_N f(gn \cdot o)\, dn$$

so

$$\hat{F}(g) = \int_N F(gn)\, dn. \tag{3}$$

PROPOSITION 5.1. *The mapping $f \to \hat{f}$ is a one-to-one linear mapping of $C^\infty_c(S)$ into $C^\infty_c(\hat{S})$.*

Proof. Consider the mapping ψ of $(K/M) \times A_\mathfrak{p}$ onto G/MN given by $\psi(kM, a) = kaMN$. Then

$$(\hat{f} \circ \psi)(kM, a) = \int_N F(kan)\, dn.$$

Now $F \in C^\infty_c(G)$ and the mapping $(k, a, n) \to kan$ is a diffeomorphism of $K \times A_\mathfrak{p} \times N$ onto G. Hence $\hat{f} \circ \psi \in C^\infty_c((K/M) \times A_\mathfrak{p})$ and since ψ is a diffeomorphism, $\hat{f} \in C^\infty_c(\hat{S})$. Next, suppose $\hat{f} \equiv 0$ and consider the function

$$f_1(g) = \int_K F(kg)\, dk,$$

where dk is the normalized Haar measure on K. Then f_1 is bi-invariant under K. Now let u be any function in $C_c(G)$, bi-invariant under K, and put

$$F_u(a) = e^{\rho(\log a)} \int_N u(an)\,dn, \qquad a \in A_{\mathfrak{p}},$$

where $\log a$ is the unique element H in $\mathfrak{h}_{\mathfrak{p}_0}$ such that $\exp H = a$. Let ν be any linear mapping of $\mathfrak{h}_{\mathfrak{p}_0}$ into $\boldsymbol{C}$ and consider the corresponding spherical function

$$\phi_\nu(x) = \int_K e^{(i\nu-\rho)(H(xk))}\,dk,$$

where for each $g \in G$, $H(g)$ is the unique element in $\mathfrak{h}_{\mathfrak{p}_0}$ such that $g \in K \exp H(g) N$. Then, for a suitable normalization of the invariant measure dx on G,

$$\int_G u(x)\,\phi_\nu(x)\,dx = \int_{A_{\mathfrak{p}}} F_u(a)\,e^{i\nu(\log a)}\,da$$

(see [11], p. 262). Now since $\hat{f} \equiv 0$ we have $F_{f_1} = 0$. Then by [13], p. 410 and p. 453 we conclude that $f_1 \equiv 0$. In particular, $F(e) = f(o) = 0$. Since $(f^{\tau(x)})^{\wedge} = (\hat{f})^{\tau(x)}$ for each $x \in G$, it follows that $f \equiv 0$ and the proposition is proved.

Remark. The image of the space $C^\infty{}_c(S)$ under the Radon transform is properly contained in $C^\infty{}_c(\hat{S})$. In fact, let ϕ be an arbitrary function in $C^\infty{}_c(A_{\mathfrak{p}})$ and consider the function $\Phi \in C^\infty{}_c(\hat{S})$ defined by $\Phi(\psi(kM, a)) = \phi(a)$ for $k \in K$, $a \in A_{\mathfrak{p}}$, where ψ is the diffeomorphism from Prop. 3.6. Suppose $f \in C^\infty{}_c(S)$ were a function such that $\hat{f} = \Phi$. Since $\Phi^{\tau(k)} = \Phi$ and $(f^{\tau(k)})^{\wedge} = \hat{f}^{\tau(k)}$ for all $k \in K$ we may, replacing f by the function $p \to \int f(k \cdot p)\,dk$, assume that f is invariant under the action of K on S. Since the function F_u above is invariant under W (Lemma 17, [11], p. 261) we deduce that the function $a \to \phi(a)e^{-\rho(\log a)}$ is invariant under W. This is impossible, ϕ being arbitrary.

We shall now require a theorem which is essentially a restatement of some results of Harish-Chandra [11], (Lemma 3, p. 247 and Theorem 1, p. 260). Subsequently a different proof was indicated by Karpelevič [16].

If ν is a linear function on $\mathfrak{h}_{\mathfrak{p}_0}$ then $e^\nu \in C^\infty(\mathfrak{h}_{\mathfrak{p}_0})$. For simplicity, the function $a \to e^{\nu(\log a)}$ on $A_{\mathfrak{p}}$ shall also be denoted e^ν.

THEOREM 5.2. *The notation being as in the preceding sections, let $\boldsymbol{D}_0(G)$ denote the set of differential operators in $\boldsymbol{D}(G)$ which are invariant under all right translations from K. Then*

(i) *For each* $D \in \boldsymbol{D}(G)$ *there exists a unique element* $D_{\mathfrak{a}} \in \boldsymbol{D}(A_{\mathfrak{p}})$ *such that*

$$D - D_{\mathfrak{a}} \in \mathfrak{n}_0 \boldsymbol{D}(G) + \boldsymbol{D}(G)\mathfrak{k}_0.$$

(ii) *If* $\phi \in C^\infty(G)$ *such that* $\phi(ngk) = \phi(g)$ *for all* $n \in N$, $g \in G$, $k \in K$ *then*

$$(D\phi)^- = D_{\mathfrak{a}}\bar{\phi}, \qquad D \in \boldsymbol{D}(G),$$

where the bar denotes restriction to $A_{\mathfrak{p}}$.

(iii) *The mapping* $D \to e^{-\rho} \circ D_{\mathfrak{a}} \circ e^{\rho}$ *is a homomorphism of* $\boldsymbol{D}_0(G)$ *onto* $I(\mathfrak{h}_{\mathfrak{p}_0})$ *and the kernel is* $\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\mathfrak{k}_0$.

In this connection we recall that a C^∞ function f on a manifold can be regarded as a differential operator $F \to fF$ and that $\circ$ denotes composition of differential operators. Also $\boldsymbol{D}(A_{\mathfrak{p}})$ is canonically isomorphic to $S(\mathfrak{h}_{\mathfrak{p}_0})$. Under this isomorphism, the unique automorphism $p \to {}'p$ of $S(\mathfrak{h}_{\mathfrak{p}_0})$ given by ${}'H = H - \rho(H)$ $(H \in \mathfrak{h}_{\mathfrak{p}_0})$ corresponds to the automorphism $D \to e^{\rho} \circ De^{-\rho}$ of $\boldsymbol{D}(A_{\mathfrak{p}})$. Statements (i) and (iii) follow immediately from the cited results of Harish-Chandra if we apply the anti-automorphism $D \to D^*$ of $\boldsymbol{D}(G)$, D^* denoting the adjoint of D. For (ii) one has to prove

$$[DT\phi](a) = [XD\phi](a) = 0$$

for $D \in \boldsymbol{D}(G)$, $T \in \mathfrak{k}_0$, $X \in \mathfrak{n}_0$, $a \in A_{\mathfrak{p}}$. But $T\phi \equiv 0$ so $DT\phi \equiv 0$; secondly

$$[XD\phi](a) = \lim_{t \to 0} 1/t([D\phi](a \exp tX) - [D\phi](a)).$$

This expression vanishes since $a \exp tXa^{-1} \in N$, whence

$$[D\phi](a \exp tX) = [(D\phi)^{L(a\exp(-tX)a^{-1})}](a) = [D\phi](a).$$

The factor algebra $\boldsymbol{D}_0(G)/\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\mathfrak{k}_0$ is canonically isomorphic to $\boldsymbol{D}(G/K)$, (see [13], p. 432). In view of (iii) and Theorem 4.1 we have two isomorphisms onto,

$$\Gamma : \boldsymbol{D}(S) \to I(\mathfrak{h}_{\mathfrak{p}_0}), \qquad \hat{\Gamma} : \boldsymbol{D}(\hat{S}) \to S(\mathfrak{h}_{\mathfrak{p}_0}).$$

THEOREM 5.3. *Let* ${}'\boldsymbol{D}(\hat{S}) = {}'\boldsymbol{D}(G/MN)$ *denote the set of* $E \in \boldsymbol{D}(\hat{S})$ *such that* ${}'\hat{\Gamma}(E)) \in I(\mathfrak{h}_{\mathfrak{p}_0})$. *Let* $D \to \hat{D}$ *denote the isomorphism of* $\boldsymbol{D}(S)$ *onto* ${}'\boldsymbol{D}(\hat{S})$ *given by*

$$'(\hat{\Gamma}(\hat{D})) = \Gamma(D), \qquad D \in \boldsymbol{D}(S).$$

Then

$$(Df)^\wedge = \hat{D}\hat{f} \qquad f \in C^\infty(S).$$

Proof. If $a \in A_{\mathfrak{p}}$, the automorphism $n \to ana^{-1}$ of N has Radon-Nikodym derivative

$$\det(\mathrm{Ad}(a) \mid \mathfrak{n}) = e^{2\rho(\log a)}$$

so

$$(4) \qquad \int_N \phi(na)\,dn = e^{2\rho(\log a)} \int_N \phi(an)\,dn, \qquad \phi \in C^\infty{}_c(A_{\mathfrak{p}}N).$$

The function $F = f \circ \pi$ and $\hat{F} = f \circ \hat{\pi}$ are related by

$$\hat{F}(g) = \int_N F(gn)\,dn$$

so

$$\hat{F}(a) = e^{-2\rho(\log a)} \int_N F(na)\,dn.$$

For $g \in G$, let $A(g)$ denote the unique element in $\mathfrak{h}_{\mathfrak{p}_0}$ such that $g = n \exp A(g) k$ for some $n \in N$ and $k \in K$. Then the function F^* given by

$$F^*(g) = e^{-2\rho(A(g))} \int_N F(ng)\,dn$$

satisfies $F^*(ngk) = F^*(g)$ for all $g \in G$, $n \in N$, $k \in K$. Let ξ denote the function $g \to e^{2\rho(A(g))}$ and as before let $F \to \bar{F}$ denote the restriction of a function on G to $A_{\mathfrak{p}}$. Then if $E \in \boldsymbol{D}(G)$ we have by Theorem 5.2 and (4)

$$(5) \qquad \int_N [EF](na)\,dn = [E(\xi F^*)](a) = [E_{\mathfrak{a}}(\bar{\xi}(\hat{F})^-)](a).$$

If $E \in \boldsymbol{D}_0(G)$ then the transform $(EF)^\wedge$ is defined and by (5)

$$(6) \qquad (EF)^\wedge(a) = \xi^{-1}(a)[E_{\mathfrak{a}}(\bar{\xi}(F^\wedge)^-)](a).$$

Let $D \in \boldsymbol{D}(S)$ be determined by $E(f \circ \pi) = (Df) \circ \pi$ for all $f \in C^\infty{}_c(S)$. Then $\Gamma(D) = e^{-\rho} \circ E_{\mathfrak{a}} \circ e^{\rho}$, by the definition of Γ. Hence, by (6),

$$(Df)^\wedge(\xi_0) = (EF)^\wedge(e) = [(e^{-\rho}\Gamma(D)e^{\rho})(\hat{F})^-](e)$$
$$= [\hat{\Gamma}(\hat{D})((\hat{F})^-)](e) = [\hat{D}\hat{f}](\xi_0),$$

the last equality following from (8) § 4. Since $(f^{\tau(g)})^\wedge = (f^\wedge)^{\tau(g)}$ for all $g \in G$ we conclude that $(Df)^\wedge(g \cdot \xi_0) = [\hat{D}\hat{f}](g \cdot \xi_0)$ and the theorem is proved.

Now let Σ denote the set of all linear functions on $\mathfrak{h}_{\mathfrak{p}_0}$ which are restrictions of some member of P_+. Let

$$\Sigma_0 = \{\lambda \in \Sigma \mid \lambda/n \notin \Sigma \text{ for all integers } n \neq 1\}.$$

Let π denote the product of all elements in Σ_0. Each element s in the Weyl

group W induces a permutation of $\Sigma \cup (-\Sigma)$ (see [13], Cor. 2.11, Ch. VII). It follows that $\pi^s = \epsilon(s)\pi$ where $\epsilon(s)$ is a real number. But $s \to \epsilon(s)$ is a homomorphism of W into $\boldsymbol{R}$ so $\epsilon(s) = \pm 1$. Consequently π^2 is a Weyl group invariant. Since B is nondegenerate on $\mathfrak{h}_{\mathfrak{p}_0}$ we identify $\mathfrak{h}_{\mathfrak{p}_0}$ with its dual and therefore $S(\mathfrak{h}_{\mathfrak{p}_0})$ with the algebra of polynomial functions on $\mathfrak{h}_{\mathfrak{p}_0}$. Since B is invariant under W, $I(\mathfrak{h}_{\mathfrak{p}_0})$ is hereby identified with the algebra of W-invariant polynomial functions on $\mathfrak{h}_{\mathfrak{p}_0}$. Let $\square$ denote the operator in $\boldsymbol{D}(S)$ which satisfies $\Gamma(\square) = \pi^2$ and let $\bar{\square}$ be any operator in $\boldsymbol{D}_0(G)$ which goes into $\square$ by the natural homomorphism μ of $\boldsymbol{D}_0(G)$ onto $\boldsymbol{D}(G/K)$ ([13], p. 390).

6. The inversion formula for complex G. Since the Radon transform $f \to \hat{f}$ is one-to-one it is a natural problem to try to find an inversion formula for it. In this section such a formula will be proved in the case when the group $I_0(S)$ is a complex semisimple Lie group.

In view of the duality between points and horocycles there is a natural dual to the transform $f \to \hat{f}$. This dual transform associates to each function $\psi \in C(\hat{S})$ a function $\check{\psi}$ on S given by

$$\check{\psi}(p) = \int_{\xi \cap p = p} \psi(\xi)\, dm(\xi), \qquad p \in S,$$

where the integral on the right is the average of ψ over the (compact) set of horocycles passing through p. Select $g \in G$ such that $g \cdot o = p$ and as before let $\xi_0 = N \cdot o$. Then

$$\check{\psi}(g \cdot o) = \int_N \psi(gk \cdot \xi_0)\, dk.$$

Let $f \in C_c(S)$ and put $I_f = (\hat{f})^\vee$. Our problem is to relate f and I_f. We have

$$I_f(p) = \int_K \left(\int_{\xi_0} f(gk \cdot s_0)\, ds_0 \right) dk = \int_K \int_N f(gkn \cdot o)\, dn\, dk.$$

For the function $F = f \circ \pi$ we put

$$I_F(g) = \int_K \int_N F(gkn)\, dn\, dk$$

and of course we have $I_F(g) = I_f(g \cdot o)$.

THEOREM 6.1. *Let S be a Riemannian globally symmetric space of the noncompact type. Assume that the group $I_0(S)$ has a complex structure. Then*

$$\square I_f = cf, \qquad f \in C^\infty{}_c(S),$$

where c is a constant $\neq 0$, independent of f.

By assumption, the Lie algebra $\mathfrak{g}_0$ has a complex structure, say J. In this case the decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + J\mathfrak{k}_0$ is a Cartan decomposition of $\mathfrak{g}_0$ so $\mathfrak{p}_0 = J\mathfrak{k}_0$. For $\mathfrak{h}_0$ we can now take the subalgebra $\mathfrak{h}_{\mathfrak{p}_0} + J\mathfrak{h}_{\mathfrak{p}_0}$ of $\mathfrak{g}_0$ so in this case $\mathfrak{h}_{\mathfrak{p}_0} = J\mathfrak{h}_{\mathfrak{k}_0}$. Now, under the scalar multiplication $(a+ib)X = aX + bJX$, $\mathfrak{g}_0$ is a Lie algebra over $\boldsymbol{C}$ and $\mathfrak{h}_0$ is a Cartan subalgebra. Let Δ' denote the set of nonzero roots of $\mathfrak{g}_0$ with respect to $\mathfrak{h}_0$. For each $\alpha \in \Delta'$, let H'_α denote the vector in $\mathfrak{h}_0$ such that $B'(H'_\alpha, H) = \alpha(H)$ $(H \in \mathfrak{h}_0)$, B' denoting the Killing form of $\mathfrak{g}_0$, considered as a Lie algebra over $\boldsymbol{C}$. Then $H'_\alpha \in \mathfrak{h}_{\mathfrak{p}_0}$ and α is real valued on $\mathfrak{h}_{\mathfrak{p}_0}$ ([13], pp. 220, 238). Fix a basis $H_1, \cdots, H_l$ of $\mathfrak{h}_{\mathfrak{p}_0}$ over $\boldsymbol{R}$ and for each root $\alpha \in \Delta'$ write $H'_\alpha = \sum_{i=1}^{l} \alpha^i H_i$. Here we define $\alpha \gg 0$ if $\alpha \neq 0$ and $\alpha^j > 0$ if j is the least index such that $\alpha^j \neq 0$. Let Q denote the set of roots $\alpha \in \Delta'$ such that $\alpha \gg 0$ and put $\rho' = \frac{1}{2} \sum_{\alpha \in Q} \alpha$. For any complex numbers $a_j = b_j + ic_j$ $(1 \leqq j \leqq l)$ we put $H_a = \sum_{j=1}^{l} b_j H_j + c_j(JH_j)$ and

$$\left(\frac{\partial}{\partial a_j} = \frac{1}{2}\frac{\partial}{\partial b_j} - i\frac{\partial}{\partial c_j}\right), \quad \frac{\partial}{\partial \bar{a}_j} = \frac{1}{2}\left(\frac{\partial}{\partial b_j} + i\frac{\partial}{\partial c_j}\right);$$

$$D_\alpha = \sum_{j=1}^{l} \alpha^j \frac{\partial}{\partial a_j}, \qquad \bar{D}_\alpha = \sum_{j=1}^{l} \alpha^j \frac{\partial}{\partial \bar{a}_j}.$$

The following formula is the principal step in the proof of the Plancherel formula for G: (see Gelfand-Neumark [6], Harish-Chandra [8], Gelfand-Graev [3])

$$(1) \qquad F(e) = c \lim_{H_a \to 0} \prod_{\alpha \in Q} D_\alpha \bar{D}_\alpha \exp(\rho'(H_a) + (\rho'(H_a))^-) \int_{K \times N} F(k \exp H_a n k^{-1})\, dn dk.$$

Here F is any function in $C^\infty{}_c(G)$ and c is a constant $\neq 0$ independent of F.

Now regarding $\mathfrak{g}_0$ as a Lie algebra over $\boldsymbol{R}$ we consider $\mathfrak{g}$, $\mathfrak{h}$, Δ as in § 3. Then it is known (see [8], p. 513) that the relation between Δ and Δ' is as follows: For each $\alpha \in \Delta'$ let α_+ and α_- denote the $\boldsymbol{C}$-linear functions on $\mathfrak{h}$ such that

$$(2) \qquad \alpha_+(H) = \alpha(H), \qquad \alpha_-(H) = (\alpha(H))^- \qquad H \in \mathfrak{h}_0$$

the bar denoting complex conjugation. Then Δ consists of all α_+, α_- as α runs through Δ'. Since each $\alpha \in \Delta'$ is real on $\mathfrak{h}_{\mathfrak{p}_0}$ the set of restrictions of elements in Δ' to $\mathfrak{h}_{\mathfrak{p}_0}$ is the same as the set of restrictions of elements in Δ to $\mathfrak{h}_{\mathfrak{p}_0}$. Consequently there is a unique Weyl chamber in $\mathfrak{h}_{\mathfrak{p}_0}$, say C, where all $\alpha \in Q$ are > 0. A root $\alpha \in \Delta'$ satisfies $\alpha \gg 0$ if and only if $\alpha > 0$ with respect to the ordering of Δ' induced by C. Now we consider C^*, Δ, P_+, P_- and ρ, defined

in § 3. Let $\alpha \in \Delta'$. Since $\mathfrak{h}_{\mathfrak{k}_0} = J\mathfrak{h}_{\mathfrak{p}_0}$ and since α is complex-linear it cannot vanish identically on $\mathfrak{h}_{\mathfrak{p}_0}$. From (2) we deduce then that $P_- = \emptyset$ and $P_+ = \{\alpha_+, \alpha_- \mid \alpha \in Q\}$. Now $H_a = H_b + H_c$ where $H_b \in \mathfrak{h}_{\mathfrak{p}_0}$, $H_c \in \mathfrak{h}_{\mathfrak{k}_0}$. Each $\alpha \in \Delta'$ is real on $\mathfrak{h}_{\mathfrak{p}_0}$ and purely imaginary on $\mathfrak{h}_{\mathfrak{k}_0}$. Hence by (2)

$$(3) \qquad \rho'(H_a) + (\rho'(H_a))^- = 2\rho'(H_b) = \rho(H_b).$$

Now suppose the function F in (1) is invariant under all right translations from K. Using (3) we see that the expression inside the braces in (1) is equal to

$$(4) \qquad \exp(\rho(H_b)) \int_{K\times N} F(k \exp H_b n)\, dk dn = e^{\rho(\log b)} \int_{K\times N} F(kbn)\, dk dn,$$

where $b = \exp H_b = \exp(\sum_{j=1}^{l} b_j H_j)$. Now, if $\phi \in C^\infty(A_{\mathfrak{p}})$ and $\Phi \in C^\infty(\boldsymbol{R}^l)$ such that

$$\phi(b) = \Phi(b_1, \cdots, b_l), \qquad b = \exp(\sum_{j=1}^{l} b_j H_j)$$

then

$$[H'_\alpha \phi](b) = \sum_{j=1}^{l} \alpha^j \frac{\partial \Phi}{\partial b_j}.$$

Hence we obtain from (1), (4), and (4) § 5

$$(5) \qquad F(e) = c_1 \lim_{b\to e} \prod_{\alpha\in Q} (H'_\alpha)^2 \{e^{-\rho(\log b)} \int_{K\times N} F(knb)\, dk\, dn\},$$

where $b \to e$ in $A_{\mathfrak{p}}$ and $c_1 = 2^{-p}c$ where p, as before, denotes the number of elements in P_+. As earlier, let B denote the Killing form of $\mathfrak{g}_0$ (as a Lie algebra over $\boldsymbol{R}$). Since B is nondegenerate on $\mathfrak{h}_{\mathfrak{p}_0} \times \mathfrak{h}_{\mathfrak{p}_0}$ we identify $\mathfrak{h}_{\mathfrak{p}_0}$ with its dual via B. Now $B(X, Y) = 2\mathrm{Re}(B'(X, Y))$ for $X, Y \in \mathfrak{g}_0$ so under the identification of $\mathfrak{h}_{\mathfrak{p}_0}$ with its dual, H'_α corresponds to $2\bar{\alpha}$ $(\alpha \in \Delta')$ where $\bar{\alpha}$ denotes the restriction of α to $\mathfrak{h}_{\mathfrak{p}_0}$. On the other hand, since α_+ and α_- have the same restriction to $\mathfrak{h}_{\mathfrak{p}_0}$ the function π^2 from § 5 coincides with $\prod_{\alpha\in Q} \bar{\alpha}^2$. This last product is therefore invariant under the Weyl group W and since B is also invariant, the same is the case for the operator $\prod_{\alpha\in Q} (H'_\alpha)^2 \in S(\mathfrak{h}_{\mathfrak{p}_0})$.

The function F_1 defined by

$$F_1(g) = \int_{K\times N} F(kng)\, dk\, dn, \qquad g \in G,$$

satisfies $F_1(ngk) = F_1(g)$ so $[DF_1](b) = [D_a\bar{F}_1](b)$ for $D \in \boldsymbol{D}(G)$ by Theorem 5.2. In particular,

$$[\bar{\square} F_1](b) = (\tfrac{1}{2})^p[(e^\rho \circ \prod_{\alpha\in Q} (H'_\alpha)^2 \circ e^{-\rho})\bar{F}_1](b).$$

Hence by (5)

$$F(e) = c \lim_{b \to e} [\tilde{\square} F_1](b). \tag{6}$$

Now let x be a fixed element in G and consider the functions $F^x = f^{\tau(x^{-1})} \circ \pi$ and F_1^x given by

$$F_1^x(g) = \int_{K \times N} F^x(kng)\, dk\, dn.$$

Then (6) implies

$$F(x) = c \lim_{b \to e} [\tilde{\square} F_1^x](b). \tag{7}$$

Also, since

$$F_1^x(g) = \int_N \left(\int_K F(xkng)\, dk \right) dn$$

we have

$$[\tilde{\square} F_1^x](g) = \int_N \left(\int_K [\tilde{\square} F](xkng)\, dk \right) dn,$$

because as g runs through a compact set in G, the integration can be taken over a compact set, independent of g. Now let $g = b$ and let $b \to e$ in $A_{\mathfrak{p}}$. Then, by (7)

$$f(x \cdot o) = F(x) = c \int_N \left(\int_K [\tilde{\square} F](xkn)\, dk \right) dn$$

which, by [13], pp. 442-443, equals

$$c \int_N \square_x \left(\int_K F(xkn)\, dk \right) dn = c \tilde{\square}_x \left(\int_{K \times N} F(xkn)\, dk\, dn \right),$$

where the subscript x denotes the argument on which $\tilde{\square}$ acts. Since the last integral is $I_F(x)$ it follows that

$$f(x \cdot o) = c[\tilde{\square} I_F](x) = c[\square I_F](x \cdot o)$$

and the theorem is proved.

Remarks. The classical Radon transform in $\boldsymbol{R}^n$ associates to each function $f \in C_0(\boldsymbol{R}^n)$ the integrals of f over hyperplanes in $\boldsymbol{R}^n$ and the problem of representing f by means of these integrals was solved by Radon [18] and John [14]. The definition of Radon transform on S is suggested by the analogy between horocycles and hyperplanes and the inversion formula in Theorem 6.1 is analogous to the inversion formula of Radon-John for $\boldsymbol{R}^n$ where n is odd. For n even the inversion formula is more complicated and involves an integral operator in place of a differential operator. Such complications are to be expected for S in the case when the group $I_0(S)$ is real.

The Radon transform for $\boldsymbol{R}^n$ was generalized in different directions to Riemannian manifolds of constant curvature by Helgason [12], Semyanistyi [21] and Gelfand-Graev-Vilenkin [5]. In the case when S is the space of positive definite Hermitian $n \times n$ matrices a formula closely related to that of Theorem 6.1 was given by Gelfand-Graev [4] and Gelfand [2].

7. Applications to invariant differential equations. In this section, the group $I_0(S)$ will be assumed complex. We shall now indicate how Theorem 6.1 can be used to reduce a differential equation

$$Du = f \tag{1}$$

where $D \in \boldsymbol{D}(S)$, $u \in C^\infty(S)$, $f \in C^\infty{}_c(S)$, to a differential equation with constant coefficients on a Euclidean space. The procedure is reminiscent of the method of plane waves for solving homogeneous hyperbolic equations with constant coefficients (see e.g. John [15]). The heuristic discussion will be followed by a concrete example, carried out rigorously.

For $x \in G$, $k \in K$, $a \in A_{\mathfrak{p}}$ we put

$$U_{k,x}(a) = \int_N u(xkan \cdot o)\, dn, \qquad F_{k,x}(a) = \int_N f(xkan \cdot o)\, dn,$$

ignoring convergence questions for the time being. Then ($A(g)$ being as in § 5) the functions

$$U_{k,x}{}^*(g) = e^{-2\rho(A(g))} \int_N u(xkng \cdot o) dn$$
$$F_{k,x}{}^*(g) = e^{-2\rho(A(g))} \int_N f(xkng \cdot o)\, dn$$

are invariant under each substitution $g \to n_1 g k_1$ and coincide with $U_{k,x}$ and $F_{k,x}$ on $A_{\mathfrak{p}}$. Using Theorem 5.2 we deduce that

$$(e^{-\rho} \circ \Gamma(D) \circ e^{\rho}) U_{k,x} = F_{k,x}. \tag{2}$$

Now by Theorem 6.1

$$f(x \cdot o) = c \tilde{\square}_x \Big(\int_K F_{k,x}(e)\, dk \Big). \tag{3}$$

This suggests the following way to write down a solution of (1): The function f being known, equation (2) is a differential equation with constant coefficients on the Euclidean space $A_{\mathfrak{p}}$. Assume now we have some "boundary conditions" such that (2) has a unique solution $U_{k,x}$ depending continuously on k and differentiably on x. Then, as suggested by (3), a solution to (1) is given by

$$u(x\cdot o) = c\bar{\square}_x\left(\int_K U_{k,x}(e)\,dk\right). \tag{4}$$

This will be verified below in a special case, where the stated assumptions are indeed satisfied.

Example: The wave equation on S. Let Δ denote the Laplace-Beltrami operator on S and let $f \in C^\infty{}_c(S)$. We shall write down a global solution of the wave equation

$$\Delta u = \frac{\partial^2 u}{\partial t^2}, \qquad (t \geqq 0) \tag{5}$$

with initial data

$$u(p,0) = 0; \quad \{\frac{\partial}{\partial t}u(p,t)\}_{t=0} = f(p) \qquad (p \in S). \tag{6}$$

Let $|\rho|^2 = B(\rho,\rho)$ (under the identification of $\mathfrak{h}_{\mathfrak{p}_0}$ with its dual). Now the Laplacian Δ on S is induced by the Casimir operator in $\boldsymbol{D}(G)$ ([13], p. 451) and we see therefore from [11], Cor. 2, p. 271 that $\Gamma(\Delta) = \Delta_A - |\rho|^2$, where Δ_A denotes the Laplacian on $A_{\mathfrak{p}}$. We are therefore led to consider the differential equation on $A_{\mathfrak{p}} \times \boldsymbol{R}^+$,

$$(\Delta_A - |\rho|^2)V^t{}_{k,x} = \frac{\partial^2}{\partial t^2}V^t{}_{k,x} \tag{7}$$

with initial data

$$V^0{}_{k,x} = 0; \quad \{\frac{\partial}{\partial t}V^t{}_{k,x}\}_{t=0} = e^\rho F_{k,x}, \tag{8}$$

where $F_{k,x}$ is the same as above. The initial value problem (7), (8) is the Cauchy problem for damped waves on the Euclidean space $A_{\mathfrak{p}}$. The solution can be found in [19], p. 88 and we obtain

$$V^t{}_{k,x}(a) = \int_{|\log b - \log a| \leqq t} (e^\rho F_{k,x})(b)\,W(t,a,b)\,db, \tag{9}$$

where

$$W(t,a,b) = 2^{-\frac{1}{2}l}(\pi\,|\rho|\,d)^{1-\frac{1}{2}l}J_{1-\frac{1}{2}l}(|\rho|\,d)$$

and $d = (t^2 - |\log a - \log b|^2)^{\frac{1}{2}}$. Here J_n refers to the usual Bessel function and as before l denotes the rank of S. The integral (9) is to be understood as its regularized value by means of analytic continuation as in [19], p. 88. As shown in [20], p. 35, $V^t{}_{k,x}$ is then a convolution (in the group $\boldsymbol{R} \times A_{\mathfrak{p}}$)

of a distribution (independent of k and x) and the function $e^{\rho}F_{k,x}$. It follows that $V^t{}_{k,x}(a)$ depends differentiably on k and x and, as suggested by (4) and the fact that $e^{\rho}(e)=1$, we consider the function

$$u(x\cdot o,t)=c\bar{\Box}_x\Big(\int_K V^t{}_{k,x}(e)\,dk\Big).$$

We shall now verify that this function is a solution of the initial value problem (5), (6). For this we need a lemma. Let $\bar{\Delta}$ be an operator in $\boldsymbol{D}_0(G)$ such that $\mu(\bar{\Delta})=\Delta$, μ being as in § 5.

LEMMA 7.1. *Let $a\in A_{\mathfrak{p}}$. Then*

$$\bar{\Delta}_x\Big(\int_K V^t{}_{k,x}(a)\,dk\Big)=(e^{-\rho}\circ\bar{\Delta}_{\mathfrak{a}}\circ e^{\rho})_a\Big(\int_K V^t{}_{k,x}(a)\,dk\Big),\tag{10}$$

where $\bar{\Delta}_{\mathfrak{a}}$ is given by Theorem 5.2.

It will first be shown that both sides of the equation (10) are solutions to the initial value problem

$$(\Delta_A-|\rho|^2)W^t{}_x=\frac{\partial^2}{\partial t^2}W^t{}_x \qquad (t\geqq 0)\tag{11}$$

$$W^0{}_x\equiv 0,\quad \Big[\frac{\partial}{\partial t}W^t{}_x\Big](a,0)=\bar{\Delta}_x\Big(\int_K (e^{\rho}F_{k,x})(a)\,dk\Big),\tag{12}$$

so the lemma will follow from the uniqueness for the Cauchy problem (11)-(12). There is no difficulty in verifying that the function

$$(t,a)\to\bar{\Delta}_x\Big(\int_K V^t{}_{k,x}(a)\,dk\Big)$$

satisfies (11) and (12) so now let us put

$$W^t{}_x(a)=(e^{-\rho}\circ\bar{\Delta}_{\mathfrak{a}}\circ e^{\rho})_a\Big(\int_K V^t{}_{k,x}(a)\,dk\Big).$$

Then $W^0{}_x\equiv 0$ and

$$\Big[\frac{\partial}{\partial t}W^t{}_x\Big](a,0)=(e^{-\rho}\circ\bar{\Delta}_{\mathfrak{a}}\circ e^{\rho})_a\Big(\int_K (e^{\rho}F_{k,x})(a)\,dk\Big).$$

In order to show that this equals the right hand side of (12) it suffices to prove that for each $D\in\boldsymbol{D}_0(G)$,

$$D_x\Big(\int_K F_{k,x}(a)\,dk\Big)=(e^{-2\rho}\circ D_{\mathfrak{a}}\circ e^{2\rho})_a\Big(\int_K F_{k,x}(a)\,dk\Big).\tag{13}$$

Since

$$\int_K F_{k,x}^*(g)\, e^{2\rho(A(g))}dk = \int_N \Big(\int_K f(xkng\cdot o)\,dk\Big)\,dn$$

we have

$$e^{2\rho(A(g))}D_x\Big(\int_K F_{k,x}^*(g)\,dk\Big) = \int_N \Big(D_x\Big(\int_K f(xkng\cdot o)\,dk\Big)\Big)\,dn,$$

which by [13], p. 443 equals

$$\int_N\Big(\int_K [\mu(D)f](xkng\cdot o)\,dk\Big)\,dn = D_g\Big(\int_N \int_K f(xkng\cdot o)\,dk\,dn\Big).$$

Putting $g = a \in A_{\mathfrak{p}}$ we have by Theorem 5.2,

$$e^{2\rho(\log a)}D_x\Big(\int_K F_{k,x}(a)\,dk\Big) = (D_{\mathfrak{a}})_a\Big(\int_N \int_K f(xkna\cdot o)\,dk\,dn\Big)$$

and now (13) follows from (4) § 5. Finally (7) shows after a straightforward computation that $W^t{}_x$ satisfies (11) so the lemma follows.

Since (6) is an immediate consequence of (3) and (8) it remains just to verify (5). Using $\square\Delta = \Delta\square$ and Lemma 7.1 we have

$$\begin{aligned}\tilde{\Delta}_x(u(x\cdot o,t)) &= c\tilde{\square}_x\Big(\tilde{\Delta}_x\Big(\int_K V^t{}_{k,x}(e)\,dk\Big)\Big)\\ &= c\tilde{\square}_x\Big(\Big[(e^{-\rho}\circ\tilde{\Delta}_{\mathfrak{a}}\circ e^{\rho})\int_K V^t{}_{k,x}dk\Big](e)\Big),\end{aligned}$$

which equals

$$c\tilde{\square}_x\Big(\int_K \frac{\partial^2}{\partial t^2}(V^t{}_{k,x}(e))\,dk\Big) \tag{14}$$

as a consequence of (7) and the relation $e^{-\rho}\circ\Delta_{\mathfrak{a}}\circ e^{\rho} = \Delta_A - |\rho|^2$ already used. In (14) we can take $\partial^2/\partial t^2$ outside the integral sign and interchange the two differential operators since they act on different arguments. Hence the expression (14) reduces to

$$\frac{\partial^2}{\partial t^2}\,(u(x\cdot o,t)).$$

This completes the verification of the solution for (5) and (6).

REFERENCES.

[1] A. Borel, "Ensembles fondamentaux pour les groupes algébriques," *Colloque sur la théorie des groupes algébriques*, Bruxelles, 1962.

[2] I. M. Gelfand, "Integral geometry and its relation to the theory of grouh representations," *Uspekhi Matematičeskih Nauk*, vol. 15 (1960), pp. 155-164; Russian Mathematical Surveys 1960, pp. 143-151.

[3] I. M. Gelfand and M. I, Graev, "Ein Analogon der Plancherelschen Formel für die klassischen Gruppen," *Trudy Moscovskogo Matematičeskogo Obščestva*, vol. 4 (1955), pp. 375-404. (German transl. in [6]).

[4] ———, "The geometry of homogeneous spaces, group representations in homogeneous spaces and questions in integral geometry related to them, I," *Trudy Moscovskogo Matematičeskogo Obščestva*, vol. 8 (1959), pp. 321-390.

[5] I. M. Gelfand, M. I. Graev, and N. Vilenkin, *Integral geometry and its relation to problems in the theory of group representations* (Generalized Functions Vol. 5) Moscov, 1962.

[6] I. M. Gelfand and M. A. Neumark, "Unitäre Darstellungen der klassischen Gruppen," *Trudy Matematičeskogo Institute im. V. A. Steklova*, vol. 36, Moscov-Leningrad, 1950. German transl. Akademie-Verlag, Berlin, 1957.

[7] R. Godement, "Groupes linéaires algébriques sur un corps parfait," *Séminaire Bourbaki*, 1960.

[8] Harish-Chandra, "The Plancherel formula for complex semisimple Lie groups," *Transactions of the American Mathematical Society*, vol. 76 (1954), pp. 485-528.

[9] ———, "On a lemma of Bruhat," *Journal de Mathématiques Pures et Appliques*, vol. 35 (1956), pp. 203-210.

[10] ———, "A formula for semisimple Lie groups," *American Journal of Mathematics*, vol. 79 (1957), pp. 733-760.

[11] ———, "Spherical functions on a semisimple Lie group I, II," *American Journal of Mathematics*, vol. 80 (1958), pp. 241-310, 553-613.

[12] S. Helgason, "Differential operators on homogeneous spaces," *Acta Mathematica*, vol. 102 (1959), pp. 239-299.

[13] ———, *Differential Geometry and Symmetric Spaces*, Academic Press, New York, 1962.

[13a] ———, "Duality and Radon transform for symmetric spaces," *Bulletin of the American Mathematical Society* (to appear).

[13b] ———, "Fundamental solutions of invariant differential operators on symmetric spaces," *Bulletin of the American Mathematical Society* (to appear).

[14] F. John, "Bestimmung einer Funktion aus ihren Integralen über gewisse Mannigfaltigkeiten," *Mathematische Annalen*, vol. 100 (1934), pp. 488-520.

[15] ———, *Plane Waves and Spherical Means Applied to Partial Differential Equations*, Interscience, New York, 1955.

[16] F. I. Karpelevič, "Orispherical radial parts of Laplace operators on symmetric spaces," *Doklady Akademii Nauk*, vol. 143 (1962), pp. 1034-1037; *Soviet Mathematics*, vol. 3 (1962), pp. 528-531.

[17] G. D. Mostow, "On maximal subgroups of real Lie groups," *Annals of Mathematics*, vol. 74 (1961), pp. 503-517.

[18] J. Radon, "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, *Berichte über die Verhandlungen der Sächsischen Akademie,* vol. 69 (1917), pp. 262-277.

[19] M. Riesz, "L'integrale de Riemann-Liouville et le problème de Cauchy," *Acta Mathematica,* vol. 81 (1949), pp. 1-223.

[20] L. Schwartz, *Théorie des Distributions II,* Hermann et Cie, Paris, 1951.

[21] E. B. Vinberg, "The Morosov-Borel theorem for real Lie groups," *Soviet Mathematics,* vol. 2 (1961), pp. 1416-1419.

EQUIVALENCE AND DECOMPOSITION OF VECTOR FIELDS ABOUT AN ELEMENTARY CRITICAL POINT.*[1]

By Kuo-Tsai Chen.

A critical point of a vector field X (i. e. an autonomous system of differential equations) is a point where X vanishes. Assume that $X = \sum a^i(x)\partial/\partial x^i$ is defined about the origin 0 of the real n-space R^n with the coordinates $x = (x^i)$. A critical point p of X will be said to be elementary if each eigenvalue of the matrix $(\partial a^i/\partial x^j)_p$ has nonvanishing real part. Our main results are summarized in the two theorems given below:

Theorem of Decomposition. *If a C^∞ vector field X has an elementary critical point at* 0, *then, in a neighborhood of* 0,

$$X = S + N$$

such that

(a) *S and N are C^∞ vector fields with $[S, N] = 0$.*

(b) *with respect to a suitable C^∞ coordinate system y,*

$$S = \sum_{i,j} c_j{}^i y^j \partial/\partial y^i$$

where the matrix $(c_j{}^i)$ is similar to a (possibly complex) diagonal matrix.

(c) *the linear part of N can be represented by a nilpotent matrix.*

The above theorem yields a nonlinear analogy of the decomposition of a linear transformation of a finite dimensional vector space into semisimple and nilpotent parts as given by the Jordan canonical form.

Theorem of Equivalence. *Let $X = \sum a^i(x)\partial/\partial x^i$ and $Y = \sum b^i(x)\partial/\partial x^i$ be two C^∞ vector fields having* 0 *as an elementary critical point. Denote by $\hat{a}^i(x)$ and $\hat{b}^i(x)$ the respective Taylor's expansions of $a^i(x)$ and $b^i(x)$ in x. Then there exists a C^∞ transformation about* 0 *which carries X to Y if and only if there exists a formal transformation which carries the formal vector field $\sum \hat{a}^i(x)\partial/\partial x^i$ to $\sum \hat{b}^i(x)\partial/\partial x^i$.*

This result reduces the problem of C^∞ equivalence of vector fields about an elementary critical point to a formal and algebraic one.

* Received June 20, 1963; revised September 27, 1963.

[1] The work has been partially supported by the National Science Foundation under Grant NSF-GP-481.

The theorems will be proved in three steps: First, a formal decomposition of X is established. In the second step, a decomposed C^∞ vector field corresponding to the formal decomposition of X is constructed. In the third step, we show that the vector field thus constructed is equivalent to X.

In §§ 2-3 we establish Lemma 3.1, which is a modification of Sternberg's wedge method for diffeomorphisms [13], [14]. Our main aim at this stage is to show Lemma 6.1. § 4 gives us a necessary tool for dealing with estimates for vector fields. With preparatory work in § 5, we obtain Lemma 6.2, which enables us to complete the proof of Lemma 6.1 through Lemma 3.1. The third step of the proof of Theorem of Decomposition is therefore finished.

The first step is given in §§ 7-8. Instead of seeking a formal transformation $\hat{\sigma}$, under which two formal vector fields become equivalent, we look for a formal vector field $\hat{Z}$ such that $\hat{\sigma}$ lies on the one-parameter group generated by $\hat{Z}$. This device yields a more "Lie-algebraic" formal treatment, which turns out to be rather helpful.

The second step (§§ 9-10) requires a further study of formal normal forms of vector fields so that the C^∞ vector field constructed is decomposed and equipped with stable and unstable manifolds. The proof for the two theorems is finally completed in § 11.

It may be said that the central problem involved in this paper is the study of local behavior of solutions of an autonomous system of differential equations about an elementary critical point. The main qualitative result in this direction seems to be the existence of stable and unstable manifolds. For references, see [6] and [12]. The same result can be obtained as a corollary of the main theorems of this paper.

Another approach is to reduce an autonomous system to a normal form. The analytic case has been investigated by H. Poincaré, G. Birkhoff, C. L. Siegel and S. Sternberg. For references, see [15]. The corresponding results for the C^∞ case turn out to be more satisfactory. Through results in diffeomorphisms, Sternberg obtained a sufficient condition for an autonomous system to be C^∞ equivalent to a linear diagonal system. See [2] and [13].

Our work is a continuation in this direction. We are able to remove the condition that the linear part of the vector field is diagonizable, which has been required in almost all previous works on normal forms under analytic or C^∞ equivalence.

It should be remarked here that M. Nagumo and K. Isé have given a sufficient condition for linearization under C^1 equivalence [9]. A recent result of P. Hartman [7] states that the 1-parameter group generated by a C^1 vector field about an elementary critical point can always be linearized under C^0 equivalence.

1. Preliminaries. Let $m=(m_1,\cdots,m_n)$ be an n-tuple of natural integers. Let $|m|=m_1+\cdots+m_n$. For any function $f(x)$, write

$$f_j=\partial f/\partial x^j$$

and

$$f_m=\partial^{|m|}f/\partial x^m$$

where $\partial^{|m|}/\partial x^m=\partial^{|m|}/(\partial x^1)^{m_1}\cdots(\partial x^n)^{m_n}$. When $|m|=0$, f_m will be taken to be f itself.

Denote by $\boldsymbol{F}(\mathfrak{D})$ the totality of functions of C^∞ in a subset $\mathfrak{D}$ of R^n. For $f\in\boldsymbol{F}(\mathfrak{D})$, we define

$$\| f \|_r=\| f \|_{r,\mathfrak{D}}=\sup\{|f_m(p)|: p\in\mathfrak{D}, |m|=r\}$$

and

$$|||f|||_r=|||f|||_{r,\mathfrak{D}}=\sup\{\| f \|_s: s=0,\cdots,r\}$$

We may verify by induction on r that

$$\| fg \|_r\leqq\sum_{i=0}^{r}C_{ri}\| f \|_i\| g \|_{r-i}, \tag{1.1}$$

where C_{ri} denotes the binomial coefficient. It follows that

$$|||fg|||_r\leqq 2^r|||f|||_r|||g|||_r. \tag{1.2}$$

For a C^∞ vector function $f=(f^1,\cdots,f^l)$ on $\mathfrak{D}$, we define

$$|||f|||_r=|||f|||_{r,\mathfrak{D}}=sup\{|||f^i|||_r: i=1,\cdots,l\}.$$

LEMMA 1.1. *For $r\geqq 0$, there exists a polynomial N_r of degree r in a single variable and with nonnegative coefficients such that, for any $f\in\boldsymbol{F}(\mathfrak{D})$ and for any C^r mapping ϕ from a subset $\mathfrak{D}'$ of R^l into $\mathfrak{D}$,*

$$|||f\circ\phi|||_r\leqq N_r(|||x\circ\phi|||_r)|||f|||_r$$

where $|||f\circ\phi|||_r=|||f\circ\phi|||_{r,\mathfrak{D}'}$ will bs understood.

Proof. The case $r=0$ is trivial. For $r>0$, we have

$$(f\circ\phi)_j=\sum_\lambda(f_\lambda\circ\phi)(x^\lambda\circ\phi)_j$$

and

$$|||(f\circ\phi)_j|||_{r-1}\leqq 2^{r-1}\sum_\lambda|||f_\lambda\circ\phi|||_{r-1}|||x\circ\phi|||_r.$$

Hence the lemma follows from the induction hypothesis.

LEMMA 1.2. *For $r\geqq 0$, there exists a polynomial M_r of degree r in*

two variables and with nonnegative coefficients such that, for any $f \in \mathbf{F}(\mathfrak{D})$ and for any two C^r mappings ϕ, ψ from $\mathfrak{D}'$ into $\mathfrak{D}$,

$$\begin{aligned}|||f\circ\psi - f\circ\phi|||_r \leqq n\,\|f\|_1\,\|x\circ\psi - x\circ\phi\|_r \\ + M_r(|||x\circ\phi|||_r, |||x\circ\psi|||_{r-1})\,|||f|||_r\,|||x\circ\psi - x\circ\phi|||_{r-1}\end{aligned}$$

provided $\mathfrak{D}$ is arcwise connected. Moreover $M_0 = 0$.

Remark. The lemma holds equally well for a vector function f.

Proof. We use induction on r. The case $r=0$ follows from the theorem of mean. For the case $r+1$,

$$(f\circ\psi - f\circ\phi)_j = \sum_\lambda [(f_\lambda\circ\psi)(x^\lambda\circ\psi - x^\lambda\circ\phi)_j + (f_\lambda\circ\psi - f_\lambda\circ\phi)(x^\lambda\circ\phi)_j]$$

and

$$\begin{aligned}(1.3)\qquad \|(f\circ\psi - f\circ\phi)_j\|_r \leqq \sum_{\lambda i} C_{ri}\{\|f_\lambda\circ\psi\|_i\,\|(x^\lambda\circ\psi - x^\lambda\circ\phi)_j\|_{r-i} \\ + \|f_\lambda\circ\psi - f_\lambda\circ\phi\|_i\,\|(x^\lambda\circ\phi)_j\|_{r-i}\}.\end{aligned}$$

Here

$$\begin{aligned}&\|f_\lambda\circ\psi\|_0 \leqq \|f\|_1,\\ &\|f_\lambda\circ\psi\|_i \leqq N_i(|||x\circ\psi|||_r)\,|||f|||_{r+1},\\ &\|(x^\lambda\circ\psi - x^\lambda\circ\phi)_j\|_{r-i} \leqq \|x\circ\psi - x\circ\phi\|_{r+1-i}.\end{aligned}$$

Moreover

$$\|f_\lambda\circ\psi - f_\lambda\circ\phi\|_i \leqq |||f_\lambda\circ\psi - f_\lambda\circ\phi|||_r,$$

for which the induction hypothesis applies. Therefore the right side of (1.3) is not greater than an expression of the type

$$n\,\|f\|_1\,\|x\circ\psi - x\circ\phi\|_{r+1} + M_{r+1}\,|||f|||_{r+1}\,|||x\circ\psi - x\circ\phi|||_r$$

where M_{r+1} may be further adjusted. Hence the lemma is proved.

2. Diffeomorphisms with stable manifolds. All diffeomorphisms and vector fields mentioned hereafter will be of C^∞ unless otherwise specified. By a local diffeomorphism about 0, we always mean one which leaves 0 fixed. Denote by $J(T)$ the jacobian of such a local diffeomorphism T at 0. If $J(T)$ has no eigenvalue of absolute value 1, then the fixed point 0 of T will be said to be elementary.

Let T be a given local diffeomorphism about the elementary fixed point 0 such that it possesses an unstable manifold V^+ and a stable manifold V^-. Suppose that V^+ and V^- happen to be two linear subspaces of R^n such that

$$R^n = V^+ \oplus V^-.$$

We may assume that the set $\{1,\cdots,n\}$ is the union of two disjoint sub-joint subsets $\mathfrak{I}_+$ and $\mathfrak{I}_-$ such that

(a) each element of $\mathfrak{I}_+$ precedes those of $\mathfrak{I}_-$,

(b) V^+ is given through the equations $x^i=0$, $i\in\mathfrak{I}_-$; and V^-, through the equations $x^i=0$, $i\in\mathfrak{I}_+$.

We may further assume that $\mathfrak{I}_+$ is not empty. Write $x^+=(x^i)_{i\in\mathfrak{I}_+}$ and $x^-=(x^i)_{i\in\mathfrak{I}_-}$ so that $x=(x^+,x^-)$.

Let $\pi_+\colon R^n\to V^+$ and $\pi_-\colon R^n\to V^-$ be the projections. Let $J_+(T)$ be the restriction of the jacobian $J(T)$ on V^+. Similarly we define $J_-(T)$.

Denote by B the quadratic form $\sum_{i\in\mathfrak{I}_+}(x^i)^2$. There exists a positive integer l such that

$$B(J_+(T)^l p)\geqq 2B(p). \qquad p\in V^+.$$

Define Q_+ to be the quadratic form such that

$$Q_+=B+B\circ J_+(T)+\cdots+B\circ J_+(T)^{l-1},$$

which is positive definite on V^+. Moreover

$$Q_+\circ J_+(T)-Q_+=B\circ J_+(T)^l-B$$

is also a positive definite quadratic form on V^+. Similarly we construct a positive definite quadratic form Q_- on V^- such that the quadratic form $Q_-\circ J_-(T)-Q_-$ is negative definite on V^-.

Define

$$Q(x)=Q_+(x^+)+Q_-(x^-).$$

Then Q is positive definite in R^n.

For $p\in R^n$, we define

$$\|x(p)\|=Q(p)^{\frac{1}{2}},$$
$$\|x(p)\|_+=Q_+(p)^{\frac{1}{2}},$$
$$\|x(p)\|_-=Q_-(p)^{\frac{1}{2}}.$$

Write, for $\rho>0$,

$$S_\rho=\{p:\|x(p)\|<\rho\},$$
$$S_\rho{}^+=\{p:\|x(p)\|_+\geqq\|x(p)\|_-\}\cap S_\rho,$$
$$S_\rho{}^-=\{p:\|x(p)\|_+\leqq\|x(p)\|_-\}\cap S_\rho,$$
$$S_\rho{}^0=\{p:\|x(p)\|_+=\|x(p)\|_-\}\cap S_\rho.$$

Note that, in $S_\rho{}^+$,

$$2\|x\|_+^2 \geqq \|x\|^2.$$

Let U be another local diffeomorphism about 0 such that $J(U) = J(T)$.

LEMMA 2.1. *There exist positive numbers ρ_0, α, β with $\alpha < \beta < 1$ such that, in $S_{\rho_0}{}^+$,*

$$\alpha\|x\|_+^2 \leqq \|x\|_+^2 - \|x\circ T^{-1}\|_+^2 \leqq \beta\|x\|_+^2, \tag{2.1}$$

$$\alpha\|x\|_+^2 \leqq \|x\circ U\|_+^2 - \|x\|_+^2 \leqq \beta\|x\|_+^2, \tag{2.2}$$

$$\alpha\|x\|_-^2 \leqq \|x\circ T^{-1}\|_-^2 - \|x\|_-^2 \leqq \beta\|x\|_-^2. \tag{2.3}$$

Proof. Since $Q_+\circ J_+(T) - Q_+$ and $Q_- - Q_-\circ J_-(T)$ are positive definite on V^+ and V^- respectively, there exist positive numbers α' and β' such that

$$\alpha'\|x\|_+^2 \leqq \|x\|_+^2 - \|x\circ J(T)^{-1}\|_+^2 \leqq \beta'\|x\|_+^2,$$
$$\alpha'\|x\|_+^2 \leqq \|x\circ J(T)\|_+^2 - \|x\|_+^2 \leqq \beta'\|x\|_+^2,$$
$$\alpha'\|x\|_-^2 \leqq \|x\circ J(T)^{-1}\|_-^2 - \|x\|_-^2 \leqq \beta'\|x\|_-^2.$$

All we need to show is that, in $S_{\rho_0}{}^+$ for some $\rho_0 > 0$,

$$\|x\circ T^{-1}\|_+^2 - \|x\circ J(T)^{-1}\|_+^2 = o(\|x\|_+^2), \tag{2.4}$$

$$\|x\circ U\|_+^2 - \|x\circ J(T)\|_+^2 = \|x\circ U\|_+^2 - \|x\circ J(U)\|_+^2 = o(\|x\|_+^2), \tag{2.5}$$

$$\|x\circ T^{-1}\|_-^2 - \|x\circ J(T)^{-1}\|_+^2 = o(\|x\|_-^2). \tag{2.6}$$

Since, in $S_{\rho_0}{}^+$, $o(\|x\|_+^2) = o(\|x\|^2)$, we have immediately (2.4) and (2.5). Observe that

$$\|x\circ T^{-1}\|_-^2 - \|x\circ J(T)^{-1}\|_-^2$$

vanishes on V^+ and has vanishing linear terms. Hence (2.6) also holds and the lemma is proved.

Remark. Since ρ_0 may be arbitrarily small, we demand that T, T^{-1} and U are defined and of C^∞ in the closure of S_{ρ_0} with $\rho_0 < 1$ and that the inequalities analogous to (2.1), (2.2) and (2.3) hold in $S_{\rho_0}{}^-$.

COROLLARY 1. *In $S_{\rho_0}{}^+$,*

$$(1-\beta)\|x\|_+^2 \leqq \|x\circ T^{-1}\|_+^2 \leqq (1-\alpha)\|x\|_+^2,$$
$$(1+\alpha)\|x\|_-^2 \leqq \|x\circ T^{-1}\|_-^2 \leqq (1+\beta)\|x\|_-^2.$$

COROLLARY 2. *If $\rho_1 = \rho_0/2$ and $p \in S_{\rho_1}{}^+$, and if l is the least positive integer such $T^{-l}p \notin S_{\rho_0}{}^+$, then*

$$\|x(T^{-l}p)\|_- > \|x(T^{-l}p)\|_+.$$

Proof. Since

$$\| x(T^{-l}p) \|_+^2 \leqq (1-\alpha)^l \| x(p) \|_+^2 < \rho_1,$$

the inequality $\| x(T^{-l}p) \|_- \leqq \| x(T^{-l}p) \|_+$ would imply

$$\| x(T^{-l}p) \|^2 < 2\rho_1^2 < \rho_0^2$$

which is absurd. Hence the corollary holds.

For $p \in S_{\rho_1}{}^+$ (or $p \in S_{\rho_1}{}^-$), define $k(p)$ to be the least natural integer such that

$$T^{-k(p)-1}p \notin S_{\rho_0}{}^+ \text{ (or } T^{k(p)+1}p \notin S_{\rho_0}{}^-).$$

When p lies in V^+ or V^-, we set $k(p) = \infty$. We leave $k(0)$ undefined.

Write $k = k(p)$. We have, in $S_\rho{}^+ - V^+$,

$$(1-\beta)^k \| x(p) \|_+^2 \leqq \| x(T^{-k}p) \|_+^2 \leqq (1-\alpha)^k \| x(p) \|_+^2,$$
$$(1+\alpha)^k \| x(p) \|_-^2 \leqq \| x(T^{-k}p) \|_-^2 \leqq (1+\beta)^k \| x(p) \|_-^2,$$

and

$$\| x(T^{-k}p) \|_- \leqq \| x(T^{-k}p) \|_+.$$

Thus we obtain

$$\| x(p) \|_-^2 / \| x(p) \|_+^2 \leqq (1-\alpha)^{k(p)} / (1+\alpha)^{k(p)}. \tag{2.7}$$

LEMMA 2.2. *There exists a positive integer k_0 and a positive number $c < 1$ such that, for $k(p) > k_0$, $p \in S_{\rho_1}{}^+$,*

$$\| x \circ T^{-1} \| \leqq c \| x \|.$$

Proof. Choose $\lambda > 0$ such that

$$c^2 = (1-\alpha) + \lambda(1+\beta) < 1.$$

It follows from (2.7) that there exists k_0 such that $k(p) > k_0$ implies

$$\| x(p) \|_-^2 / \| x(p) \|_+^2 \leqq \lambda.$$

Hence

$$\| x \circ T^{-1} \|^2 \leqq \| x \circ T^{-1} \|_+^2 + \| x \circ T^{-1} \|_-^2 \leqq (1-\alpha) \| x \|_+^2 + (1+\beta) \| x \|_-^2$$
$$\leqq ((1-\alpha) + \lambda(1+\beta)) \| x \|_+^2 \leqq c^2 \| x \|^2.$$

3. Main lemma for local diffeomorphisms. For any function f defined in S_ρ, denote by $|||f|||_{r\rho}$ the C^r norm $|||f|||_{r\rho}$ over S_ρ. If f is not of C^r at some points of S_ρ, then $|||f|||_{r\rho}$ will be understood to mean the C^r norm over S_ρ excluding those points.

LEMMA 3.1. *Let T be a local diffeomorphism about the elementary fixed point 0 such that it possesses V^+ and V^- as its unstable and stable manifolds, where V^+ and V^- are linear subspaces of R^n with $R^n = V^+ \oplus V^-$. Let U be a local diffeomorphism having contact of infinite order with T at 0. Let R^n be normed as in § 2 such that the integral valued function $k(p)$ may be defined about 0. Let there be given a local homeomorphism σ_1 about 0 having contact of infinite order with the identity mapping at 0 such that*

$$\sigma_1 p = \begin{cases} p & \text{when } \| x(p) \|_+ = \| x(p) \|_- \\ UT^{-1}p & \text{when } \| x(p) \|_+ > \| x(p) \|_-, k(p) > 0 \\ U^{-1}Tp & \text{when } \| x(p) \|_+ < \| x(p) \|_-, k(p) > 0. \end{cases}$$

It is further assumed that σ_1 is of C^∞ about 0 except at some points p with either Tp or $T^{-1}p$ lying on the cone $\| x \|_+ = \| x \|_-$. Then, given $\rho_1 > 0$, there exists $\rho_2 > 0$ satisfying the following conditions:

(a) *The local homeomorphisms σ_k from S_{ρ_2} into S_{ρ_1}, $k > 1$, given by*

$$\sigma_k p = \begin{cases} \sigma_{k-1} p & \text{when } k(p) < k-1 \\ U\sigma_{k-1}T^{-1}p & \text{when } \| x(p) \|_+ > \| x(p) \|_-, k(p) \geqq k-1 \\ U^{-1}\sigma_{k-1}Tp & \text{when } \| x(p) \|_+ < \| x(p) \|_-, k(p) \geqq k-1 \end{cases}$$

are well defined.

(b) *For $r = 0, 1, \cdots$ and for s arbitrarily large, there exist $K_s(r) > 0$ with*

$$||| x \circ \sigma_k - x |||_{r\rho} \leqq K_s(r) \rho^s$$

for $k = 1, 2, \cdots$ and $0 < \rho < \rho_2$.

(c) *$\{\sigma_k\}$ converges uniformly to a local diffeomorphism σ over some neighborhood of 0 such that*

$$U = \sigma T \sigma^{-1}$$

about 0.

Proof. We adopt the notations in § 2 including the choice of ρ_1. Each local homeomorphism σ_k may be defined individually over a sufficiently small neighborhood of 0. If σ_k is well defined in S_{ρ_2}, then it will be of C^∞ in S_{ρ_2} except at some points p with either $T^k p$ or $T^{-k} p$ lying on the cone $\| x \|_+ = \| x \|_-$.

Let k_0 and c be given as in Lemma 2.2 and let N_r, M_r be given as in Lemmas 1.1 and 1.2. Since U and T have contact of infinite order at 0, we have

$$||| x \circ (UT^{-1}) - x |||_{r\rho} \leqq Q_s(r) \rho^s, \qquad 0 < \rho < \rho_1,$$

where $Q_s(r)$ is a positive constant.

We shall prove (a) and (b) by double induction on r and k. For simplicity, we shall carry out the induction step only in $S_{\rho_2}{}^+$. First we require ρ_2 to be so small that (a) holds for $k \leqq k_0 + 1$. Write

$$N_r = N_r(|||x \circ T^{-1}|||_{r\rho_1})$$

and

$$L(r) = nN_r \| x \circ U \|_{1\rho_1}.$$

Let $l = l(r)$ be a positive integer such that $L(r)c^l < 1$. If $r \geqq 0$, there exists $K_s(r) > 0$ with

$$|||x \circ \sigma_k - x|||_{r\rho} \leqq K_s(r)\rho^s$$

for $k \leqq k_0 + 1$ and $\rho < \rho_2$. Adjust $K_s(0)$ such that

$$K_s(0) > L(0)c^s K_s(0) + Q_s(0), \qquad s \geqq l(0). \tag{3.1}$$

Write $M_r = M_r(1, K_0(r))$ and $C(r) = N_r M_r \| x \circ U \|_{r\rho_1}$. For $r > 0$, $s \geqq l$, adjust $K_s(r)$ such that

$$K_s(r) > L(r)c^s K_s(r) + C(r)c^s K_s(r-1) + Q_s(r). \tag{3.2}$$

Adjust ρ_2 such that $\rho_2 < \rho_1$ and

$$|x^i(p)| < (K_{l(0)}(0) + 1)\rho_2, \qquad i = 1, \cdots, n, \tag{3.3}$$

implies $p \in S_{\rho_1}$. Write

$$\mathfrak{D}_k = S_\rho{}^+ \cap \{p : k(p) \geqq k\}.$$

It follows from Lemma 2.2 that, for $k > k_0$ and $\rho < \rho_1$,

$$T^{-1}\mathfrak{D}_k \subset S_{c\rho}{}^+.$$

Consider now the case $(r, k+1)$ with $r \geqq 0$, $k \geqq k_0 + 1$. It is clear that, in $\mathfrak{D}_k$ with $\rho < \rho_2$,

$$\sigma_{k+1} = U\sigma_k T^{-1}$$

is well defined. Moreover

$$\begin{aligned} |||x \circ \sigma_{k+1} - x|||_{r\mathfrak{D}_k} \leqq\ & |||x \circ (U\sigma_k T^{-1}) - x \circ (UT^{-1})|||_{r\mathfrak{D}_k} \\ & + |||x \circ (UT^{-1}) - x|||_{r\mathfrak{D}_k}. \end{aligned}$$

It follows from Lemmas 1.1 and 1.2 that

$$\begin{aligned} |||x \circ (U\sigma_k T^{-1}) - x \circ (UT^{-1})|||_{r\mathfrak{D}_k} & \leqq N_r |||(x \circ U) \circ \sigma_k - x \circ U|||_{rc\rho} \\ & \leqq L(r)|||x \circ \sigma_k - x|||_{rc\rho} + C(r)|||x \circ \sigma_k - x|||_{r-1c\rho}. \end{aligned}$$

Owing to (3.1) and (3.2), we obtain

$$\begin{aligned}|||x\circ\sigma_{k+1}-x|||_{r\mathfrak{D}_s} &\leqq [L(r)K_s(r)c^s+C(r)K_s(r-1)c^s+Q_s(r)]\rho^s\\ &\leqq K_s(r)\rho^s.\end{aligned}$$

Since $\sigma_{k+1}p=\sigma_k p$ for $k(p)<k$, we conclude that (b) holds for σ_{k+1}.

For $l=l(0)$ and for $p\in S_\rho{}^+$, $\rho<\rho_2$.

$$\begin{aligned}|x^i(\sigma_{k+1}p)| &\leqq |x^i(\sigma_{k+1}p)-x^i(p)|+|x^i(p)|\\ &\leqq K_l(0)\rho^l+\rho<(K_l(0)+1)\rho_2.\end{aligned}$$

Thus, by (3.3), we have

$$\sigma_{k+1}p\in S_{\rho_1}.$$

Thus (a) also holds.

It may be easily verified that

$$\sigma_k p=\begin{cases}U^{k(p)}\sigma_1 T^{-k(p)}p & \text{for } k(p)<k\\ U^kT^{-k}p & \text{for } k(p)\geqq k.\end{cases}$$

Write $q=Tp$. Then $k(q)=k(p)+1\geqq 1$. We have, for $k(p)<k-1$,

$$\begin{aligned}\sigma_kTp=\sigma_kq &= U^{k(q)}\sigma_1T^{-k(q)}q\\ &= U^{k(p)+1}\sigma_1T^{-k(p)}p=U\sigma_kp\end{aligned}$$

and, for $k(p)\geqq k-1$,

$$\sigma_kTp=\sigma_kq=U^kT^{-k}q=U\sigma_{k-1}p.$$

Define

$$H(\sigma)=x\circ(U\sigma)-x\circ(\sigma T)$$

and $\tau_k\colon S_{\rho_2}\to R^n$ such that

$$x\circ\tau_k=x\circ\sigma_{k+1}-x\circ\sigma_k.$$

Then

$$H(\sigma_k)=\begin{cases}0 & \text{for } k(p)<k-1\\ x\circ(U\sigma_k)-x\circ(U\sigma_{k-1}) & \text{for } k(p)\geqq k-1\end{cases}$$

and

$$x\circ\tau_k=\begin{cases}0 & \text{for } k(p)<k\\ x\circ(U\sigma_kT^{-1})-x\circ\sigma_k=H(\sigma_k)\circ T^{-1} & \text{for } k(p)\geqq k.\end{cases}$$

Set

$$M^*{}_r=M_r(K_0(r),K_0(r)).$$

Then

(3.4) $$|||x\circ\tau_k|||_{r\rho}=|||x\circ\tau_k|||_{r\mathfrak{D}_s}\leqq N_r|||H(\sigma_k)|||_{rc\rho}$$

and

$$(3.5)\qquad \begin{aligned} |||H(\sigma_k)|||_{r\rho} &\leqq n\,|||x\circ U|||_{1\rho_1}\,|||x\circ\tau_{k-1}|||_{r\rho} \\ &\quad + M^{\ddagger}_r\,|||x\circ U|||_{r\rho_1}\,|||x\circ\tau_{k-1}|||_{r-1\rho} \\ &\leqq L_r\,|||x\circ\tau_{k-1}|||_{r\rho}, \end{aligned}$$

L_r being a constant independent of k. It is clear that, for ρ sufficiently small, there exists $B_s(r)>0$ such that

$$|||x\circ\tau_1|||_{r\rho} < B_s(r)\rho^s.$$

Applying (3.4) and (3.5) recursively, we obtain

$$|||x\circ\tau_k|||_{r\rho} \leqq (N_rL_rc^s)^{k-1}B_s(r)\rho^s$$

and

$$|||H(\sigma_k)|||_{r\rho} \leqq L_r(N_rL_rc^s)^{k-2}B_s(r)\rho^s.$$

We may take s such that

$$N_rL_rc^s < 1.$$

Hence σ is of C^∞ about 0 and $U=\sigma T\sigma^{-1}$ and the lemma is proved.

COROLLARY. *For T and U given in the preceding lemma, there always exists a local diffeomorphism σ about 0 such that $U=\sigma T\sigma^{-1}$ about 0.*

Proof. The existence of a local homeomorphism σ_1 as required in the preceding lemma is assured through Whitney's extension theorem [16]. Hence the corollary follows.

4. Lemmas on matrices. Denate by $\mathfrak{M}$ the totality of $n\times n$ matrices over R. In this section, we shall take R^n as the totality of $n\times 1$ matrices over R. For $H\in\mathfrak{M}$, define

$$\|H\| = \sup\{\|Hx\| : \|x\|=1, x\in R^n\}.$$

LEMMA 4.1. *Let $x(t)$ and $u(t)$ be R^n-valued functions continuous over a compactum $\mathfrak{X}$ such that $x(t)$ never vanishes. Let $M>0$. Let $H(t)$ be an $\mathfrak{M}$-valued function over X with $\|H(t)\|<M$ over $\mathfrak{X}$ such that*

$$H(t)x(t)=u(t), \qquad t\in\mathfrak{X}.$$

Then there exists an $\mathfrak{M}$-valued function $B(t)$ continuous over $\mathfrak{X}$ with $\|B(t)\|<M$ and

$$B(t)x(t)=u(t), \qquad t\in\mathfrak{X}.$$

Proof. Write $x=(x^i)$, $u=(u^i)$ and $H=(h_j^i)$. For any $t_0\in\mathfrak{X}$, we may construct

$$H(t_0,t)=(h_j^i(t_0,t))$$

continuous in a neighborhood $U(t_0)$ of t_0 such that

$$H(t_0,t)x(t)=u(t),$$

and

$$H(t_0,t_0)=H(t_0).$$

In fact, one of $x^i(t_0)$, says, $x^1(t_0)$ does not vanish, and we simply set $h_j^i(t_0,t)=h_j^i(t_0)$ for $j>1$ and

$$h_1^i(t_0,t)=x^1(t)^{-1}(u^i(t)-\sum_{l=2}^{n}h_l^i(t_0)x^l(t)).$$

Using the compactness of $\mathfrak{X}$, we may thus find a finite open covering $\mathfrak{B}_1,\cdots,\mathfrak{B}_r$ such that, for each $i=1,\cdots,r$, there is an $\mathfrak{M}$-valued function $B_i(t)$ continuous over $\mathfrak{B}_i$ with $\| B_i(t)\| < M$ and $B_i(t)x(t)=u(t)$ over $\mathfrak{B}_i$. Let $\{\phi_i(t)\}$, $i=1,\cdots,r$, be a partition of unity over $\mathfrak{X}$ such that $\phi_i(t)$ is continuous and non-negative over $\mathfrak{X}$ and vanishes outside of $\mathfrak{B}_i$. Set

$$B(t)=\sum_{i=1}^{r}\phi_i(t)B_i(t).$$

Since $\sum\phi_i(t)\equiv 1$, $B(t)$ meets our requirements. Hence the lemma is proved.

LEMMA 4.2. *Let $\mathfrak{X}$ be a closed interval. Let $x(t)$, $u(t)$, $H(t)$ be defined as in the preceding lemma, except that $x(t)$ may vanish somewhere in $\mathfrak{X}$. Then there exists an $\mathfrak{M}$-valued function $B(t)$ integrable over $\mathfrak{X}$ (in the Lebesgue sense) with $\| B(t)\| < M$ and $B(t)x=u(t)$ almost everywhere in $\mathfrak{X}$.*

Proof. Denote by $\mathfrak{X}_0$ the subset of $\mathfrak{X}$ over which $x(t)$ does not vanish. Then $\mathfrak{X}_0$ is open relative to $\mathfrak{X}$, and there exists a sequence of disjoint open intervals (a_1,b_1), (a_2,b_2), $\cdots$ such that

(a) $\mathfrak{X}_0-\bigcup_{i\geqq 1}(a_i,b_i)$ is of Lebesgue measure zero,

(b) $[a_i,b_i]\subset\mathfrak{X}_0$, $i=1,2,\cdots$.

It follows from the preceding lemma that there exists $B_i(t)$ defined and continuous over $[a_i,b_i]$ with $\| B(t)\| < M$ and $B(t)x(t)=u(t)$. Define

$$B(t)=B_i(t) \text{ for } t\in(a_i,b_i)$$

and

$$B(t)=0 \text{ for } t\in\mathfrak{X}-\bigcup(a_i,b_i).$$

It is clear that $B(t)$ has the required properties. Hences the lemma is proved.

LEMMA 4.3. *Let $x(t)$ and $v(t)$ be R^n-valued functions and $H(t)$, an $\mathfrak{M}$-valued function over the closed interval $[0,\tau]$ such that*

(a) $H(t)x(t)$ *and* $v(t)$ *are continuous,*

(b) $\| H(t) \| < M$.

If

(4.1) $$dx(t)/dt = H(t)x(t), \quad 0 \leqq t \leqq \tau,$$

holds, then

$$\| x(t) \| \leqq e^{Mt} \| x(0) \|.$$

Instead, if

(4.2) $$dx(t)/dt = H(t)x(t) + v(t), \quad 0 \leqq t \leqq \tau,$$

holds with $x(0) = 0$, *then*

$$\| x(t) \| \leqq \int_0^t e^{M(t-s)} \| v(s) \| ds.$$

Proof. Replace $H(t)$ by $B(t)$, which is integrable over $[0, \tau]$ and satisfies also the conditions (b) and (c). Then it is well known that, in the case of (4.1),

$$x(t) = U(t,s)x(s), \quad s, t \in [0, \tau],$$

where $U(t,s)$ is an $\mathfrak{M}$-valued function with $\| U(t,s) \| \leqq e^{M|t-s|}$. Thus

$$\| x(t) \| = \| U(t,0)x(0) \| \leqq e^{Mt} \| x(0) \|.$$

In the case of (4.2),

$$x(t) = \int_0^t U(t,s)v(s) ds.$$

Hence

$$\| x(t) \| \leqq \int_0^t e^{M(t-s)} \| v(s) \| ds.$$

5. Vector fields with stable manifolds. Let X be a vector field. We shall denote by $\exp tXp$ the integral path of X that passes through the point p when $t = 0$. Then $\exp X$ becomes a diffeomorphism.

Now we let X be a C^∞ vector field about the elementary critical point 0 with an unstable manifold V^+ and a stable manifold V^-. Let V^+, V^-, $\mathfrak{J}_+$, $\mathfrak{J}_-$, π_+, π_-, x^+, x^- be in the form as given in § 2. Denote by X_0 be the linear part of X. Let X_{0+} and X_{0-} be the respective restrictions of X_0 on V^+ and V^-. If $B = \sum_{i \in \mathfrak{J}_+} (x^i)^2$, then, for t sufficiently large, say $t \geqq l$,

$$B(\exp tX_{0+}p) \geqq 2B(p)$$

for any $p \in V^+$. We set

$$Q_+(p) = \int_0^l B(\exp sX_{0+}p) ds$$

which is obviously a positive definite quadratic form on V^+. Moreover

$$\begin{aligned} X_{0+}Q_+ &= [d(Q_+ \circ \exp tX_{0+})/dt]_{t=0} \\ &= [d/dt \int_t^{l+t} B \circ \exp sX_{0+} ds]_{t=0} \\ &= B \circ \exp lX_{0+} - B \geqq B. \end{aligned}$$

Consequently $X_{0+}Q_+$ is also a positive definite quadratic form on V^+. Similarly we construct a positive definite quadratic form Q_- on V^- such that $X_{0-}Q_-$ is negative definite on V^-.

Now we define the positive definite quadratic form

$$Q(x) = Q_+(x^+) + Q_-(x^-)$$

on R^n and the norms precisely in the same way as in § 2. Let Y be another C^∞ vector field about the critical point 0 such that the linear part of Y is also X_0. Both $X_0 \| x \|_+^2$ and $-X_0 \| x \|_-^2$ are positive definite on V^+ and V^- respectively. We may find constants $\beta' > \alpha' > 0$ such that

$$2\alpha' \| x \|_+^2 < X_0 \| x \|_+^2 < 2\beta' \| x \|_+^2, \tag{5.1}$$

$$2\alpha' \| x \|_-^2 < -X_0 \| x \|_-^2 < 2\beta' \| x \|_-^2. \tag{5.2}$$

Therefore, for $\rho > 0$ sufficiently small, there exist constants α, β with $\alpha > \alpha'$ and $\beta' > \beta > 0$ such that, in S_ρ^+,

$$2\beta \| x \|_+^2 \leqq X \| x \|_+^2 \leqq 2\alpha \| x \|_+^2, \tag{5.3}$$

and the same holds for Y.

LEMMA 5.1. *There exist positive constants $\rho_0 > 0$, $\alpha > 0$, $\beta > 0$ such that, in $S_{\rho_0}^+$,*

$$\alpha \| x \|_+ \leqq X \| x \|_+ \leqq \beta \| x \|_+, \tag{5.4}$$

$$\alpha \| x \|_+ \leqq Y \| x \|_+ \leqq \beta \| x \|_+, \tag{5.5}$$

$$\alpha \| x \|_- \leqq -X \| x \|_- \leqq \beta \| x \|_-. \tag{5.6}$$

Proof. The inequalities (5.4) and (5.5) follow from (5.3). The vector field $Z = X - X_0$ is also tangent to V^+ so that

$$Z \| x \|_-^2 = o(\| x \|_-^2)$$

uniformly in $S_{\rho_0}^+$ for ρ_0 sufficiently small. Hence (5.6) follows.

Now we choose $\rho_0 > 0$ with the additional property that X and Y are of C^∞ over the closure of S_{ρ_0}.

Consider the path $\exp\text{-}tXp$. As long as it stays in $S_{\rho_0}{}^+$, the preceding lemma implies that $\| x(\exp\text{-}tXp) \|_+$ is decreasing and that $\| x(\exp\text{-}tXp) \|_-$ is increasing.

Set $\rho_1 = \rho_0/2$. Let $p \in S_{\rho_1}{}^+$. Then

$$\| x(\exp\text{-}tXp) \|_- \leqq \| x(\exp\text{-}tXp) \|_+ < \rho_1,$$

and, consequently,

$$\| x(\exp\text{-}tXp) \| < 2\rho_1 = \rho_0,$$

i. e. $\exp\text{-}tXp \in S_{\rho_0}{}^+$ until the path meets $S_{\rho_0}{}^0$. For each $p \in S_{\rho_1}{}^+$, let

$$t(p) = \inf\{t : \| x(\exp\text{-}tXp) \|_+ = \| x(\exp\text{-}tXp) \|_-, t \geqq 0\}.$$

Similarly, we define $t(p) \leqq 0$ for $p \in S_\rho{}^-$, ρ sufficiently small. We may say that $t(p)$ is the time for the path $\exp\text{-}tXp$ to reach the cone $\| x \|_+ = \| x \|_-$. Since $\exp\text{-}tXp$ is a C^∞ function of p and t, it follows from the implicit function theorem that $t(p)$ is a C^∞ function of p. We may take ρ_0 sufficiently small such that $t(p)$ is of C^∞ in $S_{\rho_1} - V^+ - V^-$.

For $p \in V^+$ (or V^-), we define $t(p) = \infty$ (or $-\infty$). We leave $t(p)$ undefined at 0.

Observe that, for $p \in S_{\rho_1}{}^+$ and $0 \leqq t \leqq t(p)$, we have

$$e^{-\beta t} \| x(p) \|_+ \leqq \| x(\exp\text{-}tXp) \|_+ \leqq e^{-\alpha t} \| x(p) \|_+ \tag{5.7}$$

and therefore

$$\| x(\exp\text{-}tXp) \| \leqq \sqrt{2} e^{-\alpha t} \| x(p) \|_+ \leqq 2e^{-\alpha t} \| x(p) \|. \tag{5.8}$$

As a matter of fact, (5.7) and (5.8) hold for any p and t as long as the path $\exp\text{-}sXp$, $0 \leqq s \leqq t$, lies entirely in $S_{\rho_0}{}^+$.

For any $\gamma > 0$, define $t(p,\gamma) = t(p)$ or $\gamma t(p)/|t(p)|$ according as $|t(p)| \leqq \gamma$ or $|t(p)| > \gamma$. The function $t(p,\gamma)$ is continuous over $S_{\rho_1} - V^+ - V^-$ and is of C^∞ over the same region except at those points p with $|t(p)| = \gamma$.

6. Main lemma for vector fields. For any function f and any mapping σ, we write as usual

$$\sigma^* f = f \circ \sigma$$

when the composition at the right hand side is properly defined. If σ is a local diffeomorphism and if X is a vector field, we define

$$\operatorname{Adj} \sigma X = \sigma^{*-1} X \sigma^*.$$

Geometrically, the diffeomorphism σ carries the vector field X to the vector field $\operatorname{Adj} \sigma X$. For detail, see [4].

The purpose of this section is to establish the next result:

LEMMA 6.1. *Let $X = \sum a^i(x)\partial/\partial x^i$ be a vector field defined about* 0 *and having* 0 *as an elementary critical point. Let R^n be the direct sum of two linear subspaces V^+ and V^- which are the respective stable and unstable manifolds of X about* 0. *Then, for any vector field $Y = \sum b^i(x)\partial/\partial x^i$ having contact of infinite order with X at* 0, *there exists a local diffeomorphism σ about* 0 *such that*

$$\operatorname{Adj} \sigma X = Y$$

in a neighborhood of 0.

Let X be in the form as given in the preceding section. We are going to construct a diffeomorphism σ such that it leaves the cone $\| x \|_+ = \| x \|_-$ pointwise fixed and carries X to Y. Since any flow of X about 0 either lies in $V^+ \cup V^-$ or passes through the cone, the image under σ for any point about 0 not lying in $V_+ \cup V^-$ is well determined. However it is a complicated matter to show that σ is actually of C^∞ about 0. The geometric notion of such a diffeomorphism σ was used by Birkhoff and Bamforth [1] in the case $n = 3$.

Let M and A_r be positive numbers such that, in S_{ρ_0},

$$\| (a_j{}^i(x)) \| < M,$$
$$\| (b_j{}^i(x)) \| < M,$$
$$\| v(x) \| \leqq A_r \| x \|^r, \qquad r = 1, 2, \cdots$$

where $v(x) = (v^i(x)) = (b^i(x) - a^i(x))$. Define, for $\gamma \geqq 0$,

$$\sigma_\gamma p = (\exp t(p,\gamma) Y)(\exp - t(p,\gamma) X) p,$$

when p is not the origin, and

$$\sigma_\gamma 0 = 0.$$

LEMMA 6.2. *There exists ρ_2, $0 < \rho_2 < \rho_1$, such that*

(a) *for any $p \in S_{\rho_2}$ and $\gamma > 0$, the path $\exp tY \exp - t(p,\gamma) Xp$, t running from* 0 *to $t(p,\gamma)$, lies entirely in S_{ρ_0}.*

(b) *there exist $K_r > 0$, $r = 1, 2, \cdots$, with*

$$\| x(\sigma_\gamma p) - x(p) \| \leqq K_r \| x(p) \|^r$$

for $p \in S_{\rho_2}$ and $\gamma \geqq 0$.

Proof. Set

$$\xi(t) = x(\exp tYq) - x(\exp tXq).$$

If both paths $\exp tYq$ and $\exp tXq$, $0 \leqq t \leqq \tau$, lie in S_{ρ_0}, then

$$(6.1)\qquad d\xi/dt = [b(\xi + x(\exp tXq)) - b(v(\exp tXq))] + v(x(\exp tXq)),$$

with $\xi(0) = 0$. The theorem of the mean yields

$$b^i(\xi + x(\exp tXq)) - b^i(x(\exp tXq)) = \sum b_j{}^i(\eta_t)\xi^i,$$

when η_t belongs to the convex neighborhood S_{ρ_0}. Now (6.1) takes the form of (4.2), and it follows that

$$\begin{aligned}\|\xi(t)\| &\leqq \int_0^t e^{M(t-s)}\|v(x(\exp sXq))\|\,ds \\ &\leqq \int_0^t e^{M(t-s)}A_r\|x(\exp sXq)\|^r\,ds.\end{aligned}$$

Let $p \in S_{\rho_1}{}^+$ and $q = \exp - t(p,\gamma)Xp$. For simplicity, write $\gamma' = t(p,\gamma)$. Then the path $\exp tXq = \exp - (\gamma' - t)Xp$, $0 \leqq t \leqq \gamma'$, lies in $S_{\rho_0}{}^+$. As long as $\exp tYq$, $0 \leqq t \leqq \gamma'$ also stays in $S_{\rho_0}{}^+$, we shall have from (5.8),

$$\begin{aligned}(6.2)\qquad \|\xi(t)\| &\leqq A_r\int_0^t e^{M(t-s)}\|x(\exp - (\gamma' - s)Xp)\|^r\,ds \\ &\leqq A_r\int_0^t e^{M(t-s)}(2e^{-\alpha(\gamma'-s)}\|x(p)\|)^r\,ds \\ &\leqq 2^rA_r(-M + r\alpha)^{-1}e^{-r\alpha(\gamma'-t)}\|x(p)\|^r,\end{aligned}$$

provided $r > M/\alpha$. Set

$$K_r = 2^rA_r(-M + r\alpha)^{-1}$$

for $r > M/\alpha$. Choose ρ_2, $0 < \rho_2 < \rho_1/2$, such that for some $r_0 > M/\alpha$,

$$K_{r_0}(\rho_2)^{r_0} + \rho_1 < 3\rho_0/4.$$

Then, for any $p \in S_{\rho_2}{}^+$, the path $\exp - tXp$, $0 \leqq t \leqq t(p)$, lies in $S_{\rho_1}{}^+$. Consequently, the path $\exp tXq$, $0 \leqq t \leqq \gamma'$, stays in $S_{\rho_1}{}^+$. We assert that the path $\exp tYq$, $0 \leqq t \leqq \gamma'$, stays in $S_{3\rho_0/4}$. In fact, if this is not true, then there exists γ'', that is the least positive number such that

$$\|x(\exp\gamma''Yq)\| = 3\rho_0/4.$$

It follows from (6.2) that

$$\begin{aligned}\|x(\exp\gamma''Yq)\| &\leqq \|\xi(\gamma'')\| + \|x(\exp\gamma''Xq)\| \\ &\leqq K_{r_0}\rho_2{}^{r_0} + \rho_1 < 3\rho_0/4,\end{aligned}$$

which is absurd. The part (a) of the proposition is now established. For $r > M/\alpha$, we have

$$\| x(\sigma_\gamma p) - x(p) \| = \| x(\exp \gamma' Yq) - x(\exp \gamma' Xq) \|$$
$$\leqq K_r \| x(p) \|^r.$$

Since (b) holds for r sufficiently large, it holds for all $r = 0, 1, 2, \cdots$. Hence the lemma is proved.

Proof of Lemma 6.1. Write $T = \exp X$ and $U = \exp Y$. Then T has V^+ and V^- as its respective stable and unstable manifolds. According to Proposition 7.1, T and U have contact of infinite order at 0. If $k(p)$ is defined as in § 2, then $k(p)$ is the greatest integer $\leqq t(p)$. Lemma 6.2 assures that σ_1 has contact of infinite order with the identity mapping at 0. It is then evident that, with the local homeomorphism σ_1 thus defined, Lemma 3.1 may be applied. Moreover, each local homeomorphism σ_k, $k > 1$, given in the lemma coincides with that defined in this section. It remains to verify that $Y = \operatorname{Adj} \sigma X$ or, equivalently,

$$\exp tY\sigma p = \sigma \exp tXp \tag{6.3}$$

for p about 0 and $|t|$ small. Note that, for p about 0 and $p \notin V^+ \cup V^-$,

$$\sigma p = (\exp t(p) Y)(\exp - t(p) X) p,$$

and (6.3) holds.

For p about 0 and $p \in V^+$ (or V^-), we have

$$\sigma p = \lim_{k \to \infty} U^k T^{-k} p.$$

For t near zero,

$$\| x(\exp tY\sigma_k p) - x(\sigma_k \exp tXp) \|$$
$$= \| x(U^k \sigma_t \exp - (k-t)Xp) - x(U^k \exp - (k-t)Xp) \|,$$

by repeated application of the theorem of mean,

$$\leqq M \| x(U^{k-1}\sigma_t \exp - (k-t)Xp) - x(U^k \exp - (k-t)Xp) \|$$
$$\leqq \cdots \leqq M^k \| x(\sigma_t \exp - (k-t)Xp) - x(\exp - (k-t)Xp) \|,$$

by Lemma 6.2,

$$\leqq M^k K_r \| x(\exp - (k-t)Xp) \|^r$$
$$\leqq M^k K_r (2e^{-\alpha(k-t)} \| x(p) \|)^r$$

which tends to zero as $k \to \infty$ provided a suitably large r is used. Hence the lemma is proved.

7. Formalization. Again, we are going to work in the n-dimensional space R^n with a given system of coordinate (x^i), whose origin is 0.

Denote by $\boldsymbol{F}$ the totality of (the germs of) functions of C^∞ in a neighborhood of 0, by $\boldsymbol{A}$ the totality (of the germs) of vector fields of C^∞ about 0 and vanishing at 0; and by $\boldsymbol{G}$ the totality of (the germs of) the local homeomorphisms of C^∞ in a neighborhood of 0 and leaving 0 fixed. Then $\boldsymbol{F}$ is an algebra over the field R of the real numbers, $\boldsymbol{A}$ is a Lie algebra over R, and $\boldsymbol{G}$ is a group.

Denote by $\mathfrak{F}$ the algebra of the formal power series in $x = (x^i)$ of the form $\sum_{|m| \geqq 0} a_m x^m$, $a_m \in R$; by $\mathfrak{A}$ the Lie algebra of the formal vector fields of the form

$$\sum_{i=1}^{n} \hat{a}^i(x) \partial/\partial x^i,$$

where each $\hat{a}^i(x) \in \mathfrak{F}$ has a vanishing constant term; and by $\mathfrak{G}$ the group of the formal transformations $(\hat{h}'(x), \cdots, \hat{h}^n(x))$, $\hat{h}^i \in \mathfrak{F}$, such that $\hat{h}^i(0) = 0$ and the Jacobian $\det(\partial \hat{h}^i/\partial x^j)$ has a nonvanishing constant term.

For $\hat{f} \in \mathfrak{F}$, $\hat{X} \in \mathfrak{A}$, $\hat{\sigma} \in \mathfrak{G}$, the meaning of $\hat{X}\hat{f}$ and that of $\hat{\sigma}^*\hat{f}$ are clear. We define

$$\operatorname{Adj} \hat{\sigma}\hat{X} = \sum \hat{b}^i \partial/\partial x^i,$$

where

$$\hat{b}^i = (\hat{\sigma}^{-1})^* \hat{X} \hat{\sigma}^* x^i.$$

Define $\Theta_F\colon \boldsymbol{F} \to \mathfrak{F}$ such that, for $f \in \boldsymbol{F}$, $\Theta_F f$ is the Taylor's expansion of f about 0. Define $\Theta_A\colon \boldsymbol{A} \to \mathfrak{A}$ such that, if $X = \sum a^i(x)\partial/\partial x^i$, then

$$\Theta_A X = \sum (\Theta_F a^i) \partial/\partial x^i,$$

and define $\Theta_G\colon \boldsymbol{G} \to \mathfrak{G}$ such that, for $\sigma \in \boldsymbol{G}$,

$$\Theta_G \sigma = (\Theta_F\, \sigma^* x^1, \cdots, \Theta_F \sigma^* x^n).$$

We shall not distinguish Θ_F, Θ_A, Θ_G and shall denote all of them simply by Θ.

It is known that there always exists a function $f \in \boldsymbol{F}$ having a given Taylor's expansion about 0, i. e. Θ_F is onto. So are Θ_A and Θ_G. It is straightforward to verify that Θ_F, Θ_A, Θ_G are respectively algebra, Lie algebra and group homomorphisms. Moreover

$$\Theta(Xf) = (\Theta X)(\Theta f),$$
$$\Theta(\sigma^* f) = (\Theta\sigma)^*(\Theta f),$$
$$\Theta(\operatorname{Adj} \sigma X) = \operatorname{Adj}(\Theta\sigma)(\Theta X).$$

For $X \in \boldsymbol{A}$, $(\exp tX)0 \equiv 0$ for any value of t. It follows that $\exp tX$ always exists and belongs to $\boldsymbol{G}$.

PROPOSITION 7.1. *If* $\circledcirc X = \circledcirc Y$, $X, Y \in \boldsymbol{A}$, *then, for any value of* $t \in R$,

$$\circledcirc \exp tX = \circledcirc \exp tY.$$

Proof. Write $X = \sum a^i(x)\partial/\partial x^i$ and $Y = \sum b^i(x)\partial/\partial x^i$. Set $\sigma_t = \exp tX$ and $\tau_t = \exp tY$. We are going to show that, for any value of t,

$$(x^i \circ \sigma_t)_m(0) = (x^i \circ \tau_t)_m(0).$$

For $|m| = 0$, this is evidently true. Now, by induction on $|m|$, we have, for $r = |m| > 0$,

$$\begin{aligned} d[(x^i \circ \sigma_t)_m(0)]/dt &= [\partial^r(\partial x^i \circ \sigma_t/\partial t)/\partial x^m]_{x=0} \\ &= [\partial^r a^i(x \circ \sigma_t)/\partial x^m]_{x=0} \\ &= \sum_{\lambda=0}^{n} a_\lambda{}^i(0)\,(x^\lambda \circ \sigma_t)_m(0) + P^i, \qquad i = 1, \cdots, n. \end{aligned}$$

where P^i may be expressed as a polynomial in $a_{m'}{}^i(0)$, $|m'| \leqq |m|$, and $(x^\lambda \circ \sigma_t)_{m'}(0)$, $\lambda = 1, \cdots, n$, $|m'| < |m|$. Then, by induction hypothesis, $(x^i \circ \tau_t)_m(0)$, $i = 1, \cdots, n$, satisfy the same systems of linear differential equations with the same initial conditions at $t = 0$. Hence the proposition is proved.

Definition 7.1. For $\hat{X} \in \mathfrak{A}$, define

$$\exp X = \circledcirc \exp X,$$

where X is any element of $\boldsymbol{A}$ with $\circledcirc X = X$.

If $\hat{f}(t)$ is an $\mathfrak{F}$-valued function of a real variable t, $\partial f(t)/\partial t$ is understood to be obtained through term by term differentiation with respect to t. If $\hat{X}(t) = \sum \hat{a}^i(t)\partial/\partial x^i$ is an $\mathfrak{A}$-valued function of t, we define

$$\partial X(t)/\partial t = \sum (\partial \hat{a}^i(t)/\partial t)\partial/\partial x^i.$$

If $f(x,t)$ is of C^∞ in x and t for (x,t) in a neighborhood of $(0, t_0)$, then

$$\circledcirc(\partial f(x,t)/\partial t) = \partial \circledcirc f(x,t)/\partial t$$

in a neighborhood of t_0. An analogous assertion also holds for $\hat{X}(t)$.

PROPOSITION 7.2. *If* $\hat{X}, \hat{Y} \in \mathfrak{A}$ *and if* $\hat{\sigma}_t = \exp t\hat{X}$, *then*

$$\partial \operatorname{Adj} \hat{\sigma}_t \hat{Y}/\partial t = [\operatorname{Adj} \hat{\sigma}_t \hat{Y}, \hat{X}].$$

Proof. Choose $X, Y \in \boldsymbol{A}$ such that $\circledcirc X = \hat{X}$ and $\circledcirc Y = \hat{Y}$. Write

$\sigma_t = \exp tX$. It is known (see, for example, Theorem 2.2, [4]) that in a neighborhood of 0,

$$\partial \operatorname{Adj} \sigma_t Y x^i / \partial t = [\operatorname{Adj} \sigma_t Y, X] x^i.$$

Thus

$$\mathfrak{S}(\partial \operatorname{Adj} \sigma_t Y x^i / \partial t) = \mathfrak{S}[\operatorname{Adj} \sigma_t Y, X] x^i = [\operatorname{Adj} \hat{\sigma}_t \hat{Y}, \hat{X}] x^i.$$

The left hand side is equal to

$$\partial(\mathfrak{S} \operatorname{Adj} \sigma_t Y x^i)/\partial t = \partial \operatorname{Adj} \hat{\sigma}_t \hat{Y} x^i / \partial t.$$

Hence the proposition holds.

Denote by $\mathfrak{F}^{(r)}$ the totality of homogeneous polynomials of degree r in x. Then the algebra $\mathfrak{F}$ has a graded structure, which is called a formal algebra in [3]. We define $\mathfrak{A}^{(r)}$ to be the totality of $X \in \mathfrak{A}$ such that $\hat{X} = \sum \hat{a}^i(x) \partial/\partial x^i$ with each $\hat{a}^i(x) \in \mathfrak{F}^{(r+1)}$. Then $\mathfrak{A}$ has a graded structure with

$$[\mathfrak{A}^{(r)}, \mathfrak{A}^{(s)}] \subset \mathfrak{A}^{(r+s)}.$$

Let

$$\pi_r : \mathfrak{A} \to \mathfrak{A}^{(r)}, \qquad r = 0, 1, \cdots,$$

be the projection. Now $\hat{Z} = \operatorname{Adj} \hat{\sigma}_t \hat{Y}$ is the solution of the formal differential equation

$$d\hat{Z}/dt = [\hat{Z}, \hat{X}]$$

in the formal module $\mathfrak{A}$ with the initial condition $\hat{Z}(0) = \hat{Y}$. If $\pi_0 \hat{X} = 0$, then

$$\begin{aligned} \hat{Z}(t) &= \hat{Y} + t[\hat{Z}, \hat{X}] + t^2[[\hat{Z}, \hat{X}], \hat{X}]/2! + \cdots \\ &= (\exp t \operatorname{ad} \hat{X}) \hat{Y}, \end{aligned}$$

where $\operatorname{ad} \hat{X}$ denotes the linear transformation $\hat{Y} \to [\hat{Y}, \hat{X}]$. (See p. 115, [3].)

We summarize the above discussion in the following statement:

PROPOSITION 7.3. *If* $\hat{X}, \hat{Y} \in \mathfrak{A}$ *and* $\pi_0 \hat{X} = 0$, *then*

$$(\operatorname{Adj} \exp \hat{X}) \hat{Y} = (\exp \operatorname{ad} \hat{X}) \hat{Y}.$$

Definition 7.2. We say that $X, Y \in \boldsymbol{A}$ are equivalent if there exists $\sigma \in \boldsymbol{G}$ such that $Y = \operatorname{Adj} \sigma X$.

Similarly, we say that $\hat{X}, \hat{Y} \in \mathfrak{A}$ are equivalent if there exists $\hat{\sigma} \in \mathfrak{G}$ such that $\hat{Y} = \operatorname{Adj} \hat{\sigma} \hat{X}$.

8. Formal Lie algebra. Let $\mathfrak{L} = \mathfrak{L}^{(0)} + \mathfrak{L}^{(1)} + \cdots$ be a formal module over a field K with the projections $\pi_r : \mathfrak{L} \to \mathfrak{L}^{(r)}$, $r = 0, 1, \cdots$. If $\mathfrak{L}$ is furthermore a Lie algebra over K such that for any $X, Y \in \mathfrak{L}$,

$$\pi_r[X, Y] = \sum_{i=0}^{r} [\pi_i X, \pi_{r-i} Y],$$

then $\mathfrak{L}$ is called a formal Lie algebra over K.

For $X \in \mathfrak{L}$, let $\mathrm{ad}\cdot X$ be the endomorphism of $\mathfrak{L}$ such that

$$(\mathrm{ad}\, X) Y = [Y, X].$$

Definition 8.1. We shall say that $X, Y \in \mathfrak{L}$ are (formally) conjugate, if there exists $Z \in \mathfrak{L}$ with $\pi_0 Z = 0$ such that

$$Y = (\exp \mathrm{ad}\, Z) X.$$

Obviously $\mathfrak{A}$ is a formal Lie algebra over R. If $\hat{X}, \hat{Y} \in \mathfrak{A}$ are conjugate, then, according to Proposition 7.3, they are equivalent.

For any $X \in \mathfrak{L}^{(0)}$, denote by

$$\mathrm{ad}_r X \colon \mathfrak{L}^{(r)} \to \mathfrak{L}^{(r)}$$

the restriction of $\mathrm{ad}\, X$ on $\mathfrak{L}^{(r)}$. The kernel and the image of $\mathrm{ad}_r X$ will have the usual notations $\ker \mathrm{ad}_r X$ and $\mathrm{im}\, \mathrm{ad}_r X$ respectively.

PROPOSITION 8.1. *If, for $X \in \mathfrak{L}$, there exist subspaces $\mathfrak{H}^{(r)} \subset \mathfrak{L}^{(r)}$ such that $\mathfrak{L}^{(r)}$ is the sum (not necessarily direct) of $\mathfrak{H}^{(r)}$ and* $\mathrm{im}\, \mathrm{ad}_r \pi_0 X$, *then X is conjugate to some $Y \in \mathfrak{L}$ such that*

(a) $\pi_0 Y = \pi_0 X$

(b) $\pi_r Y \in \mathfrak{H}^{(r)}$, $r > 0$.

Proof. We are going to construct Z such that

$$Y = (\exp \mathrm{ad}\, Z) X$$

satisfies the requirements. First, we set $\pi_0 Z = 0$. Choose $\pi_1 Z$ such that

$$\pi_1 Y = \pi_1 X + [\pi_0 X, \pi_1 Z] = \pi_1 X - (\mathrm{ad} \pi_0 X) \pi_1 Z \in \mathfrak{H}^{(1)}.$$

In general,

$$\begin{aligned} \pi_r Y &= \pi_r X + [\pi_0 X, \pi_r Z] \\ &\quad + \sum_{i=1}^{r-1} (\mathrm{ad}\, \pi_{r-i} Z)(\pi_i X) \\ &\quad + 1/2! \sum_{i+j+k=r} (\mathrm{ad}\, \pi_i Z)(\mathrm{ad}\, \pi_j Z)(\pi_k X) + \cdots \\ &= \pi_r X - (\mathrm{ad}_r \pi_0 X) \pi_r Z + R_r \end{aligned}$$

where R_r is determined by $\pi_0 X_1, \cdots, \pi_{r-1} X$, $\pi_1 Z, \cdots, \pi_{r-1} Z$. It is then clear that there exists $\pi_r Z$ such that $\pi_r Y \in \mathfrak{H}^{(r)}$. Hence the proposition is proved.

THEOREM 8.1. *If, for* $X \in \mathfrak{L}$, $\pi_0 X = S + N$ *satisfies the conditions:*

(a) $S \in \mathfrak{L}^{(0)}$, $N \in \mathfrak{L}^{(0)}$,

(b) $[S, N] = 0$,

(c) $\mathrm{ad}_r S$ *is semi-simple and* $\mathrm{ad}_r N$ *is nilpotent*, $r = 1, 2, \cdots$,
then X *is conjugate to some* $S + N'$ *such that* $\pi_0 N' = N$ *and* $[S, N'] = 0$.

Proof. Since $\mathrm{ad}_r S$ and $\mathrm{ad}_r N$ commute, $\mathrm{ad}_r N$ leaves both $\ker \mathrm{ad}_r S$ and $\mathrm{im}\, \mathrm{ad}_r S$ invariant, and so does $\mathrm{ad}_r(S + N)$. Moreover, $\mathrm{ad}_r S$ is an automorphism on $\mathrm{im}\, \mathrm{ad}_r S$.

If $x \in \mathrm{im}\, \mathrm{ad}_r S$, $x \neq 0$, and $\mathrm{ad}_r(S + N)x = 0$, then

$$\mathrm{ad}_r N x = -\mathrm{ad}_r S x.$$

Let s be the least positive integer such that $(\mathrm{ad}_r N)^s x = 0$. Then

$$(\mathrm{ad}_r N)^s x = -(\mathrm{ad}_r S)(\mathrm{ad}_r N)^{s-1} x \neq 0,$$

which is absurd. Therefore, $\mathrm{ad}_r(S + N)$ is an automorphism on $\mathrm{im}\, \mathrm{ad}_r S$ so that

$$\mathrm{im}\, \mathrm{ad}_r S \subset \mathrm{im}\, \mathrm{ad}_r(S + N). \tag{8.1}$$

We may choose $\mathfrak{H}^{(r)} = \ker \mathrm{ad}_r S$, where $\mathfrak{H}^{(r)}$ is as given in Proposition 8.1. Set $N' = Y - S$. Thus the theorem is proved.

9. Rational n-tuples. Denote by R_0^n the vector space of the n-tuples of rational numbers. Any element of R_0^n can be written as

$$e = (m_1(e), \cdots, m_n(e)).$$

If J is a subset of $\{1, \cdots, n\}$, the totality of e with $m_j(e) = 0$, $j \in J$, will be called a face of R_0^n. It will be said to be proper when J is not empty. For any subspace $\mathfrak{U}$ of R_0^n, denote by $\mathfrak{U}^+$ the subset of $\mathfrak{U}$ consisting of all e with $m_i(e) \geqq 0$, $i = 1, \cdots, n$.

PROPOSITION 9.1. *For any subspace* $\mathfrak{U}$ *of* R_0^n, *there exist* $u_1, \cdots, u_\lambda \in \mathfrak{U}^+$ *such that every element of* $\mathfrak{U}^+$ *can be expressed as*

$$e = \alpha_1 u_1 + \cdots + \alpha_\lambda u_\lambda \tag{9.1}$$

where the α*'s are non-negative rational numbers.*

Proof. Let $\mathfrak{B}_1, \cdots, \mathfrak{B}_\lambda$ be the faces such that each $\mathfrak{B}_i$ contains at least one nonzero element, say, $u_i \in \mathfrak{U}^+$. We are going to show by induction on n that the proposition holds for the u's so chosen. The case $n = 0$ and that of $\lambda = 0$ are trivial. Assume $n > 0$ and $\lambda > 0$. Let e be any element of $\mathfrak{U}^+$.

Let α be the least non-negative number such that $e-\alpha u_1$ belongs to the intersection of $\mathfrak{U}^+$ and a proper face. The induction hypothesis now applies. Hence the proposition holds for e.

Denote by $\mathfrak{U}(k)$ the totality of the elements u of $\mathfrak{U}$ with $m_i(u)\geqq 0$ for $i\neq k$ and $m_k(u)=-1$.

PROPOSITION 9.2. *Given k, $1\leqq k\leqq n$, there exist $v_1,\cdots,v_\mu\in\mathfrak{U}(k)$ such that any $d\in\mathfrak{U}(k)$ can be expressed as*

$$d=\beta_1 v_1+\cdots+\beta_\mu v_\mu+e \tag{9.2}$$

where $e\in\mathfrak{U}^+$ with $m_k(e)=0$ and the β's are non-negative rational numbers whose sum is equal to 1.

Proof. The case $n=0$ is trivial. We proceed by induction on n. If $\mathfrak{U}(k)$ consists of one element only, then the proposition is obvious. If $\mathfrak{U}^+$ contains a nonzero element e with $m_k(e)=0$, then, for any element d of $\mathfrak{U}(k)$, there is a non-negative rational number β such that $d-\beta e$ lies in the intersection of $\mathfrak{U}(k)$ and a proper face, provided $n>1$. Then the induction hypothesis may take over. The separated case of $n=1$ is also clear.

Now assume that $\mathfrak{U}^+$ possesses no nonzero element u with $m_k(u)=0$. Let d' be an element of $\mathfrak{U}(k)$, which is distinct from d. For some i,j, we have

$$m_i(d)<m_i(d')$$

and

$$m_j(d)>m_j(d'),$$

for otherwise one of $d-d'$ and $d'-d$ has to belong to $\mathfrak{U}^+$. Set

$$d(t)=(1-t)d+td'.$$

Then, for some $t_0\leqq 0$ and $t_1\geqq 1$, both $d(t_0)$ and $d(t_1)$ belong to $\mathfrak{U}(k)$ as well as a proper face. Since the proposition holds for $d(t_0)$ and $d(t_1)$, it also holds for d and d'. Hence the proposition is proved.

10. Construction of a decomposed vector field. Our aim in this section is to construct a vector field as specified in the next theorem:

THEOREM 10.1. *Let $S=\sum_{i,j}s_j{}^i x^j\partial/\partial x^i$ and $\hat{N}\in\mathfrak{A}$ be such that*

$$[\hat{N},\circledast S]=0.$$

If the $n\times n$ matrix $(s_j{}^i)$ is semisimple, i.e., similar to a complex diagonal matrix diagonal $(\lambda_1,\cdots,\lambda_n)$, then there exists N in A such that $\circledast N=\hat{N}$ and

$$[N, S] = 0$$

First, we extend the formal Lie algebra $\mathfrak{A}$ to the formal Lie algebra $\tilde{\mathfrak{A}}$ by extending the ground field R to that of the complex numbers. By a linear change of the coordinate system, we may obtain a system of coordinates (z^i), $i = 1, \cdots, n$, such that

$$S = \sum_i \lambda_i z^i \partial/\partial z^i. \tag{10.1}$$

Since S is real, we may further demand that the coordinates $z^1, \cdots, z^n$ occur in conjugate pairs.

Let $\tilde{\mathfrak{B}}$ be the centralizer of $\circledS S$ in the Lie algebra $\mathfrak{A}$ i.e. the totality of $\hat{X} \in \mathfrak{A}$ with $[\hat{X}, \circledS S] = 0$. It is easy to see that $\tilde{\mathfrak{B}}$ is also a formal Lie algebra.

Denote by $\mathfrak{U}$ the subspace of $R_0{}^n$ that consists of all u with $\sum m_i(u)\lambda_i = 0$. We say that $u \in \mathfrak{U}$ is integral if all $m_i(u)$ are integral. Denote by $\mathfrak{U}\langle k\rangle$ the totality of $u \in \mathfrak{U}$ which are integral with $m_k(u) \geqq -1$ and $m_i(u) \geqq 0$, $i \neq k$. In short $\mathfrak{U}\langle k\rangle$ is the totality of the integral elements in $\mathfrak{U}(k) \cup \mathfrak{U}^+$.

PROPOSITION 10.1. *The totality of $z^{m(d)} z^k \partial/\partial z^k$, $d \in U\langle k\rangle$, $|m(d)| = r$, $k = 1, \cdots, n$, is a base of $\mathfrak{B}^{(r)}$, $r = 0, 1, \cdots$.*

Proof. The proposition follows immediately from the fact that

$$[z^m z^k \partial/\partial z^k, S] = -\sum m_i \lambda_i z^m z^k \partial/\partial z^k.$$

Let $u_1, \cdots, u_\lambda \in \mathfrak{U}^+$ be given as in Proposition 9.1. For a given k, let $v_1, \cdots, v_\mu \in \mathfrak{U}(k)$ be given as in Proposition 9.2.

Any $d \in \mathfrak{U}\langle k\rangle$ can be given in the form

$$d = \alpha_1 u_1 + \cdots + \alpha_\lambda u_\lambda + \beta_1 v_1 + \cdots + \beta_\mu v_\mu \tag{10.2}$$

for some nonnegative rational numbers α's and β's, where either

$$\beta_1 + \cdots + \beta_\mu = 1 \text{ or } \beta_1 = \cdots = \beta_\mu = 0.$$

Let

$$\gamma^{-1}/4 = \max\{|m(u_1)|, \cdots, |m(u_\lambda)|\}$$

and

$$\delta = \max\{|m(v_1)|, \cdots, |m(v_\mu)|\}.$$

Write any $d \in \mathfrak{U} < k >$ in the form of (10.2). Then

$$|m(d)| = \alpha_1 |m(u_1)| + \cdots + \alpha_\lambda |m(u_\lambda)| + \beta_1 |m(v_1)| + \cdots + \beta_\mu |m(v_\mu)|,$$

and, for $|m(d)| > 2\delta$,

$$\gamma^{-1}\sum \alpha_i/4 \geqq \sum \alpha_i \,|\, m(u_i) \,|\, \geqq \,|\, m(d) \,|\, - \delta > \,|\, m(d) \,|/2$$

so that

$$\sum \alpha_i > 2\gamma \,|\, m(d) \,|. \tag{10.3}$$

Each of $u_1, \cdots, u_\lambda$ may be replaced by a positive integral multiple of itself. Therefore, we may assume that $u_1, \cdots, u_\lambda$ are all integral. Set $\phi_i = z^{m(u_i)}$ and

$$\Phi = \phi_1 \bar{\phi}_1 + \cdots + \phi_\lambda \bar{\phi}_\lambda,$$

which is real valued and is defined over R^n.

Let $\chi(t)$ be a real valued C^∞ function on the real line such that $\chi(t) = 1$ for $|\, t \,| \leqq 1/2$, $\chi(t) = 0$ for $|\, t \,| \geqq 1$, and $0 \leqq \chi(t) \leqq 1$ for $1/2 < |\, t \,| < 1$.

Proof of Theorem 10.1. We shall assume that $\lambda > 0$ so that Φ exists. Otherwise each $\mathfrak{U}\langle k \rangle$ consists of a finite number of elements so that $\tilde{\mathfrak{B}}^{(r)} = \{0\}$, and thus $\pi_r \hat{N} = 0$ for r sufficiently large. The construction becomes trivial by simply setting

$$N = \sum_{r=0}^{\infty} \pi_r \hat{N}.$$

Now write

$$\pi_r \hat{N} = \sum b_r{}^i(z) z^i \partial/\partial z^i. \tag{10.4}$$

Since $\pi_r \hat{N} \in \tilde{\mathfrak{B}}^{(r)}$, we have

$$b_r{}^i(z) = \sum b^i(d) z^{m(d)}$$

summing over all $d \in \mathfrak{U}\langle i \rangle$ with $|\, m(d) \,| = r$, where each $b^i(d)$ is a complex number. Write $c_r = \sum |\, b^i(d) \,|$ summing over $i = 1, \cdots, n$ and all $d \in \mathfrak{U}\langle i \rangle$ with $|\, m(d) \,| = r$. Set

$$h^k(z) = \sum_{r=0}^{\infty} b_r{}^k(z) z^k \chi(r!\, c_r \Phi).$$

When $\Phi > 0$, the right hand side is a finite sum, and $h^k(z)$ is of C^∞ over the region: $\Phi > 0$. We are going to show that $h^k(z)$ is of C^∞ at any point p with $\Phi(p) = 0$. For this purpose, it suffices to prove that, given any positive integer s, there exists a positive integer r_0 such that

$$h_{r_0}{}^k(z) = \sum_{r=r_0}^{\infty} b_r{}^k(z) z^k \chi(r!\, c_r \Phi) \tag{10.5}$$

$$= 0(|\, x - x(p) \,|^s),$$

where $|\, x \,| = (\sum |\, x^i \,|^2)^{\frac{1}{2}}$.

Take a neighborhood $\mathfrak{N}$ of p such that, in $\mathfrak{N}$, $|\, \Phi \,| < 1$ and $|\, \operatorname{grad} \Phi \,|$ has an upper bound L. Then

$$|\Phi| = |\Phi - \Phi(p)| \leqq L\,|x - x(p)|.$$

Choose r_0 such that (a) $2\gamma r_0 > n$ and (b) $\gamma r_0 > s+2$.

It follows from (10.3) that if, for $d \in \mathfrak{U}\langle k\rangle$, $|m(d)| \geqq r_0$, d is expressed in the form of (10.2), then we have $\sum \alpha_i > n$, which implies that $\alpha_i > 1$ for some i. Since $\phi_i(p) = z^{m(u_i)}(p) = 0$, there exists j with $z^j(p) = 0$ and $m_j(u_i) \geqq 1$. Consequently

$$m_j(d) \geqq \alpha_i m_j(u_i) > 1.$$

which means that $z^{m(d)}z^k$ has z^j as a factor. Therefore we conclude that, for $d \in \mathfrak{U}\langle k\rangle$ and $|m(d)| \geqq r_0$, the function $z^{m(d)}z^k$ vanishes at p.

On the other hand, owing to $|m(\sum \beta_i v_i)| < \delta$, we may choose the neighborhood $\mathfrak{N}$ such that $|z^{m(\Sigma\beta_i v_i)}z^k|$ has an upper bound K in $\mathfrak{N}$ for any choice of $d \in \mathfrak{U}\langle k\rangle$. Then, for $|m(d)| \geqq r_0$, we have in $\mathfrak{N}$

$$\begin{aligned}|z^{m(d)}z^k| &= |\phi_1{}^{\alpha_1}\cdots\phi_\lambda{}^{\alpha_\lambda} z^{m(\Sigma\beta_i v_i)}z^k| \\ &\leqq K\Phi^{\Sigma\alpha_i/2} \leqq K\Phi^{\gamma|m(d)|}.\end{aligned}$$

If $r \geqq r_0$, and $r!c_r\Phi \leqq 1$, then

$$\begin{aligned}|b_r{}^k(z)z^k\chi(r!c_r\Phi)| &\leqq c_r K\Phi^{\gamma r} \leqq K\Phi^{\gamma r-1}/r! \\ &\leqq K\Phi^{s+1}/r! \leqq KL^{s+1}\,\|x - x(p)\|^{s+1}/r!\end{aligned}$$

When $r!c_r\Phi > 1$, then

$$|b_r{}^k(z)z^k\chi(r!c_r\Phi)| = 0.$$

With this remark, we have shown (10.5).

We now define

$$N = \sum h^i(z)\,\partial/\partial z^i.$$

Since $\pi_r\hat{N} = \sum b_r{}^i(z)z^i\partial/\partial z^i$ is real, for any $f \in \boldsymbol{F}$,

$$Nf = \sum_{r=0}^{\infty}(\pi_r\hat{N}f)\chi(r!c_r\Phi)$$

is real valued, i.e. $N \in \boldsymbol{A}$. Evidently $\circledast N = \hat{N}$.

Now

$$S\chi(r!c_r\Phi) = r!c_r\chi'(r!c_r\Phi)S\Phi = 0.$$

When $\Phi > 0$,

$$N = \sum \chi(r!c_r\Phi)\pi_r N,$$

the sum being finite, and

$$[N, S] = \sum \chi(r!c_r\Phi)\,[\pi_r\hat{N}, S] = 0.$$

The dimension of the algebraic manifold $\Phi = 0$ in R^n is less than n. Therefore $[N, S]$ vanishes throughout R^n by continuity. Thus the proof is completed.

11. Proofs of main theorems. We are now ready to prove the main results stated in the introductory part of this paper.

Proof of Theorem of Decomposition. It suffices to show that X is equivalent to $S + N$. Write $\hat{X} = \circledast X$ and $\pi_0\hat{X} = \hat{S} + \hat{N}'$, where $\hat{S}$ is semisimple and $\hat{N}'$ is nilpotent. Observe that $\mathrm{ad}_r \hat{S}$ is semisimple when it operates on $\tilde{\mathfrak{A}}^{(r)}$ and therefore on $\mathfrak{A}^{(r)}$. We assert that, in the same sense, $\mathrm{ad}_r \hat{N}'$ is nilpotent.

In fact, for a given r, there exists a positive integer l such that

$$\hat{N}'^l\mathfrak{Z}^{(0)} = \cdots = \hat{N}'^l\mathfrak{Z}^{(r+1)} = \{0\}.$$

For any $\hat{Z} \in \mathfrak{A}^{(r)}$, we then have

$$((\mathrm{ad}_r \hat{N}')^{2l}\hat{Z})z^i = (\hat{N}'^{2l}\hat{Z} - C_{2l,1}\hat{N}'^{2l-1}\hat{Z}\hat{N}' + \cdots + C_{2l,2l}\hat{Z}\hat{N}'^{2l})z^i = 0.$$

Therefore, by Theorem 8.1, $\hat{X}$ is conjugate, in $\mathfrak{A}$, to $\hat{S} + \hat{N}$ with $\pi_0\hat{N} = \hat{N}'$ and $[\hat{N}, \hat{S}] = 0$. It follows that X is equivalent to some $X' \in A$ such that $\circledast X' = \hat{S} + \hat{N}$. For simplicity we assume that X is already in the form of X'.

Let $\mathfrak{J}_+$ ($\mathfrak{J}_-$) be the subset of $\{1, \cdots, n\}$ that consists of all i with $\mathrm{Re}\,\lambda_i > 0$ ($\mathrm{Re}\,\lambda_i < 0$). Denote by V^+ (V^-) the subspace of R^n such that $z^i = 0, i \in \mathfrak{J}_-$ ($i \in \mathfrak{J}_+$). If $d \in \mathfrak{U}\langle k\rangle$ and $k \in \mathfrak{J}_+$ ($k \in \mathfrak{J}_-$), then $m_j(d) > 0$, for some $j \in \mathfrak{J}_+$ ($j \in \mathfrak{J}_-$), otherwise

$$\sum_{j\in\mathfrak{J}_+} m_j(d)\,\mathrm{Re}\,\lambda_j = \sum m_j(d)\,\mathrm{Re}\,\lambda_j + \sum_{j\in\mathfrak{J}_-} m_j(d)\,\mathrm{Re}\,\lambda_j$$

would not vanish. Therefore, if $b_r^k(z)$ is as given in (10.4), then $b_r^k(z)z^k$ vanishes on V^-. Let S and N be constructed as in Theorem 10.1. Then N and thus $S + N$ leave V^+ and V^- invariant. By Lemma 6.1, X is equivalent to $S + N$. Hence the theorem is proved.

Proof of Theorem of Equivalence. According to the preceding theorem, X is equivalent to $X' = S + N$, which leaves V^+ and V^- invariant.

Now $\circledast X'$ and $\circledast Y$ are equivalent in $\mathfrak{A}$, i.e. there exists $\hat{\sigma} \in G$ such that

$$\mathrm{Adj}\,\hat{\sigma}\circledast Y = \circledast X'.$$

Take $\sigma \in G$ such that $\circledast\sigma = \hat{\sigma}$. Write $Y' = \mathrm{Adj}\,\sigma Y$. Then $\circledast Y' = \circledast X'$. Apply again Lemma 6.1. We conclude that X and Y are equivalent.

COROLLARY 1. *If X is any vector field of C^∞ about 0 and having 0 as*

a critical point, then there exists a C^∞ stable manifold and a C^∞ unstable manifold about 0, *and the sum of the dimension of the two manifolds is* n.

COROLLARY 2. *Let $X = \sum a^i(x)\partial/\partial x^i$ be of C^∞ about* 0 *and possess an elementary critical point at* 0. *If the zeroes of the characteristic polynomial of the matrix $(a_j^i(0))$, $\lambda_1, \cdots, \lambda_n$, satisfy the inequalities*

$$\sum m_i\lambda_i \neq \lambda_j, \qquad j = 1, \cdots, n, \tag{11.1}$$

for any natural integers $m_1, \cdots, m_n$, with $|m| \geqq 2$, then X is equivalent to its linear part $X_0 = \sum a_j^i(0)x^j\partial/\partial x^i$.

Proof. Let $\mathcal{S}$ and $\hat{N}$ be given as in the proof of Theorem of Decomposition. It follows from Proposition 10.1 that (11.1) implies $\tilde{\mathfrak{B}}^{(r)} = 0$ for $r > 0$. This forces $\hat{N} = \hat{N}'$, and $\otimes X$ is now equivalent to $\otimes X_0$ in $\mathfrak{A}$. Hence X is equivalent to X_0 in A.

RUTGERS—THE STATE UNIVERSITY.

REFERENCES.

[1] G. D. Birkhoff and F. R. Bamforth, "Divergent series and singular points of ordinary differential equations," *Transactions of the American Mathematical Society*, vol. 32 (1930), pp. 114-146.

[2] F. Bruhat, "Travaux de Sternberg," *Séminaire Bourbaki*, 13 (1960/61), Exposé 217

[3] K. T. Chen, "Formal differential equations," *Annals of Mathematics*, vol. 73 (1961), pp. 110-133.

[4] ———, "Decomposition of differential equations," *Mathematische Annalen*, vol. 146 (1962), pp. 263-278.

[5] ———, "Equivalence and decomposition of local vector fields," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 49 (1963), pp. 740-741.

[6] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[7] P. Hartman, "On the local linearization of differential equations," *Proceedings of the American Mathematical Society*, vol. 14 (1963), pp. 568-573.

[8] J. Moser, "The analytic invariants of an area preserving mapping near a hyperbolic fixed point," *Communications on Pure and Applied Mathematics*, vol. 9 (1956), pp. 673-692.

[9] M. Nagumo and K. Isé, "On the normal forms of differential equations in the neighborhood of an equilibrium point," *Osaka Mathematical Journal*, vol. 9 (1957), pp. 221-234.

[10] H. Poincaré, *Oeuvres*, vol. I, Paris, 1928.

[11] C. L. Siegel, *Vorlesungen über Himmelsmechanik*, Springer, Berlin, 1956.

[12] S. Smale, " Stable manifolds for differential equations and diffeomorphisms,"*Annali della Scuola Normale Superiore di Pisa*, vol. 17 (1963), pp. 97-116.

[13] S. Sternberg, " On the structure of local homeomorphism," *American Journal of Mathematics*, vol. 80 (1958), pp. 623-631.

[14] ———, " The structure of local homeomorphisms III," *American Journal of Mathematics*, vol. 81 (1959), pp. 578-604.

[15] ———, Infinite Lie groups and the formal aspects of dynamical systems," *Journal of Mathematics and Mechanics*, vol. 10 (1961), pp. 451-474.

[16] H. Whitney, "Analytic extensions of differentiable functions defined in closed sets." *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 63-89.

ON THE REAL TRACES OF ANALYTIC VARIETIES.*[1]

By Wei-Liang Chow.

To B. L. van der Waerden on his 60th birthday

1. Let V be a complex-analytic variety of complex dimension n which is the complexification of its real trace V^0, so that V^0 is a real-analytic variety of real dimension n, and let X be a complex-analytic subvariety of complex dimension r in V such that $X^0 = X \cap V^0$ has the real dimension r (and hence $\bar{X} = X$). If we denote by $\phi(X)$ and $\phi(V^0)$ the homology classes mod 2 determined by X and V^0 respectively as topological cycles mod 2 in V, then their intersection class $\phi(X)\cdot\phi(V^0)$, with due regard for their supports, can be considered as an element in $H_{2r-n}(V^0, \mathbf{Z}_2)$. On the other hand, X^0 is also a topological cycle mod 2 in V^0 and hence determines an element $\psi(X^0)$ in $H_r(V^0, \mathbf{Z}_2)$. It is natural to ask about the relation between these two elements $\phi(X)\cdot\phi(V^0)$ and $\psi(X^0)$ in $H_*(V^0, \mathbf{Z}_2)$, since roughly speaking they represent respectively the algebraic intersection and the geometric intersection of X and V^0. We shall give here an answer to this question for the case where V and X are algebraic. Our result is surprisingly simple: We have the relation

$$\phi(X)\cdot\phi(V^0) = \psi(X^0)^2 \tag{1}$$

in the sense of numerical equivalence mod 2 for algebraic classes in $H_*(V^0, \mathbf{Z}_2)$. In case V^0 has the property that every homology class mod 2 is representable by an algebraic cycle, as for example in the case of a real Grassmannian, then the relation (1) holds in the sense of homology mod 2.

We can probably best illustrate our result by an interesting application, which was in fact the original motivation of our present inquiry. Let V and V^0 be respectively the complex and real Grassmannians of all s-spaces in the complex and real spaces of dimension m; if we consider V and V^0 as the respective classifying spaces of the complex and real s-vector bundles, then the universal dual Chern classes $c_1, \cdots, c_s$, and the universal dual Whitney classes $w_1, \cdots, w_s$, are represented by certain Schubert varieties $C_1, \cdots, C_s$ and their real traces $C_1{}^0, \cdots, C_s{}^0$ respectively. The precise sense in which the Whitney classes can be considered as the "real traces" of the Chern classes

* Received March 20, 1963.

[1] This work was partially supported by a research grant of the National Science Foundation.

has been made clear by Borel and Haefliger [1] by means of a mapping ρ of the complex-analytic cycles in V into $H_*(V^0, \boldsymbol{Z}_2)$, so that one has the relation $\psi(C_j^0) = \rho(C_j)$. Now, if we choose the dimension m so that V^0 is orientable and hence the Poincaré duality (over $\boldsymbol{Z}$) holds in V^0, then the universal dual Pontrjagin classes p_j are given by $\theta(C_j)\cdot\theta(V^0)$, $j=1,\cdots,s$, where $\theta(C_j)$ and $\theta(V^0)$ denote the homology classes over $\boldsymbol{Z}$ determined by C_j and V^0 respectively as cycles over $\boldsymbol{Z}$ in V, and the intersection classes $\theta(C_j)\cdot\theta(V^0)$ are to be considered as elements in $H_*(V^0, \boldsymbol{Z})$. Since the elements $\phi(C_j)\cdot\phi(V^0)$ can be identified with the reduction mod 2 of the elements $\theta(C_j)\cdot\theta(V^0)$, our result then yields the following well-known relations between the Pontrjagin classes and the Whitney classes (or rather their duals):

$$p_j = \omega_j^2 \mod 2.$$

A topological proof of these relations using transgressive elements can be found in [2], Appendix II.

2. Our proof of (1) is based on a theorem (Theorem 1) which gives a relation between the intersection of a cycle X with its conjugate and its intersection with V^0. Although in this paper we shall need this theorem only for the case of algebraic cycles in an algebraic variety, we shall prove it here for the general case, both on account of its own interest as well as for the sake of possible generalization of our result.

We begin by making a few comments on the intersection theory for an analytic variety, either complex or real, and the relations between the two cases. It is possible to develop an intersection theory for complex-analytic cycles in a complex-analytic variety along lines roughly similar to the now classical intersection theory for algebraic cycles in an algebraic variety. Such an intersection theory, just as in the algebraic case, is essentially unique, and it is compatible with the topological homology theory in the sense that the mapping ϕ (or θ) which assign to each complex-analytic cycle in V the homology class determined by it in $H_*(V, \boldsymbol{Z}_2)$ (or $H_*(V, \boldsymbol{Z})$) is multiplicative, i.e., $\phi(X)\cdot\phi(Y) = \phi(X\cdot Y)$ whenever $X\cdot Y$ is defined. Recently, Borel and Haefliger [1], on the basis of a homology theory with supports previously developed by Borel and Moore, were able to obtain by topological methods an intersection theory for complex-analytic cycles in a complex-analytic variety. Since in this theory the intersection multiplicities are defined by means of homology classes with supports, there is no question here of compatibility and the mapping ϕ is so to speak "built-in." For our present purpose it is immaterial which theory is used; the only property we need here is the criterion

for multiplicity by means of small deformations contained in the following lemma.

Before stating our lemma, we add a few words in explanation of a notation used in it. Let N be a neighborhood of the origin in the complex n-space and let $e = (e_1, \cdots, e_n)$ be a point in N; then we denote by T_e the translation in N defined by the mapping $z \to z + e$. As a transformation in N, T_e is of course only defined in $N \cap T_{-e}N$ and maps it onto $N \cap T_e N$; however, given any neighborhood N' of the origin such that $\bar{N}' \subset N$, we have the relation $N' \subset T_e N$ for all sufficiently small e. If X is any cycle in N, then we shall denote by $T_e X$ the cycle in $N \cap T_e X$ which is the image under T_e of the restriction $X \cap T_{-e}N$ of X in $N \cap T_{-e}N$. Thus $T_e X$ is strictly speaking not a cycle in N, but for any N' such that $\bar{N}' \subset N$, the restriction of $T_e X$ in N' is a cycle for all sufficiently small e. If N is a neighborhood of a point p in a complex-analytic variety V such that there exists a system of local coordinates in V at p which is valid in N, and if X is any cycle in N, then we can extend the above definitions of T_e and $T_e X$ to this more general case by means of the local coordinate system; and finally, if X is any cycle in V, we shall for simplicity write $T_e X$ for $T_e(X \cap N)$.

Lemma 1. *Let V be a complex-analytic variety of dimension n, and let X and Y be complex-analytic prime cycles of dimension r and $n - r$ respectively in V which intersect properly at a point p; let N be a neighborhood of p in V such that X and Y intersect only at p in N and such that there exists a system of local coordinates in V at p which is valid in N. Then there is a neighborhood N' of p in N such that for almost all points e in a sufficiently small neighborhood of p, the number of intersections of $T_e X$ with Y in N' is exactly equal to $i(X \cdot Y, p)$.*

Remark. The expression "almost all" means here "all except for a proper complex-analytic subset."

Proof. The property to be proved being local, we can without any loss of generality take V to be the coordinate neighborhood N, or equivalently take $V = N$ to be a neighborhood of the origin p in the complex n-space with the coordinates $z = (z_1, \cdots, z_n)$; furthermore, we can assume that both X and Y are locally irreducible at p, for otherwise we can treat each irreducible component separately, replacing N by a smaller neighborhood if necessary (so small that any two different local components of X or Y at p represent different subvarieties in N). Now, just as in the abstract algebraic intersection theory, we can introduce the product cycle $X \times Y$ in the product variety

$V \times V$, and show that the number of intersection points of X and T_eY in N' is the same as the number of intersection points of $X \times Y$ with $T_e\Delta$ in $N' \times N'$, where Δ is the diagonal of $V \times V$ and T_e as applied to Δ is the local translation induced in $N \times N$ by the application of the original T_e to the second factor alone. It follows then that our lemma can be reduced to the special case where Y is the linear variety in N defined by the equations $z_1 = \cdots = z_r = 0$. The condition that X and Y intersect properly at p implies that the defining ideal of X (in the power series ring of n variables $z_1, \cdots, z_n$) is regular with respect to the given coordinate system, with $z_1, \cdots, z_r$ as independent variables. We have therefore, after a suitable linear transformation of the coordinates $z_{r+1}, \cdots, z_n$ if necessary, the well-known "parametric" representation of X by equations of the form $P(z_{r+1}) = 0$; $z_i = f_i/D$, $i = r+2, \cdots, n$, where $P(z_{r+1})$ is an irreducible "distinguished polynomial" in z_{r+1} with coefficients which are power series in $z_1, \cdots, z_r$, the function D is the discriminant of the polynomial P, and the functions $f_{r+2}, \cdots, f_n$ are polynomials in z_{r+1} with coefficients which are power series in $z_1, \cdots, z_r$. This representation is valid in a sufficiently small polycylindrical neighborhood $N' = \{z \mid |z_i| < d_i\}$ of p, and is to be understood in the sense that the points in $X \cap N'$ consist precisely of all solutions of these equations in N' for which the systems of values of $z_1, \cdots, z_r$ are not zeroes of D, as well as all limits of these solutions as the set of values of $z_1, \cdots, z_r$ approaches a zero of D in N'. (We observe that this implies in particular that $X \cap N'$ has no limiting points on the boundary of N' outside of the subset satisfying the conditions $z_i = d_i$ for $i = 1, \cdots, r$.) Therefore, if t is the degree of the polynomial $P(z_{r+1})$, for any set of values of $z_1, \cdots, z_r$ with $|z_i| < d_i$ which is not a zero of D, the t distinct zeroes $P(z_{r+1})$ will give rise in this representation to t distinct points in $X \cap N'$, and these are the only points in $X \cap N'$ having the given set of values for the coordinates $z_1, \cdots, z_r$. Now, let $N'' \subset N'$ be a neighborhood of p such that for every point e in N'' we have the relation $N' \subset T_eN$ (we can evidently assume that $\bar{N}' \subset N$), then for every e in N'' the restriction of T_eY to N' is the variety defined by the equations $z_1 = e_1, \cdots, z_r = e_r$. It follows that for any such point e for which the subsystem $e_1, \cdots, e_r$ is not a zero of D, the intersection points of X with T_eY in N' are precisely the t points obtained from the above representation of X by setting $z_1 = e_1, \cdots, z_r = e_r$. Our lemma $i(X \cdot Y, p) = t$ follows then from [1], 4.10, by observing that T_eY intersects X transversally in N' and that the intersection of the set $\{z \mid |z_i| \leqq e_i, i = 1, \cdots, r\}$ with $X \cap N'$ is compact.

3. In case of a real-analytic variety V^0 the situation is somewhat similar if one takes everything over $\boldsymbol{Z}_2$ instead of $\boldsymbol{Z}$. Here we have a mapping ψ which assigns to each real-analytic cycle in V^0 the homology class determined by it in $H_*(V^0, \boldsymbol{Z}_2)$, the reason for taking $\boldsymbol{Z}_2$ being the fact that the real-analytic varieties are not always orientable. However, there is a difficulty due to the fact that in the real-analytic case the sheaf of ideals of a subvariety is not necessarily coherent, in contrast to what is known in the complex-analytic case. This leads to the unfortunate situation, first pointed out by H. Cartan, that it is not always possible to resolve globally a real-analytic subvariety into irreducible components in a reasonable manner; and since the analytic cycles are defined in terms of irreducible subvarieties, this makes it difficult to give a meaningful definition of the analytic cycles. One could remedy this situation by restricting oneself to the "coherent" subvarieties, i. e. subvarieties whose sheaves of ideals are coherent, but this would have the drawback that even some real-algebraic subvarieties defined by single real-algebraic equations would be excluded. A satisfactory compromise seems to be the notion of a C-analytic subvariety, first introduced by H. Cartan [4], which roughly speaking is a real-analytic subvariety which has a global complexification. As has been shown by Whitney and Bruhat [5], the C-analytic subvarieties have many of the "good" properties of the complex-analytic subvarieties; in particular, such a subvariety admits a unique decomposition into C-irreducible components, whereby a C-analytic subvariety is said to be C-irreducible if it is not the union of two distinct C-analytic subvarieties.

Although the way seems to be thus opened for an intersection theory for real-analytic cycles in a real-analytic variety, there does not seem to be at this moment such a theory which runs more or less parallel to the complex-analytic case and is independent of it. However, as Borel and Haefliger have observed [1], it is possible to derive an intersection theory over $\boldsymbol{Z}_2$ for real-analytic cycles from the complex-analytic case by using the complexifications, with the result that the mapping ψ mentioned above is also multiplicative. To explain this briefly, let V be the complexification of V^0 (the existence of this complexification is assured if we assume that V^0 is paracompact); then a mapping α can be defined which assigns to every complex-analytic cycle X (over $\boldsymbol{Z}$ or $\boldsymbol{Z}_2$) of complex dimension r in V a real-analytic cycle X^0 over $\boldsymbol{Z}_2$ of real dimension r in V^0, and the intersection theory in V^0 is precisely so defined as to make the mapping α multiplicative for the subset of all complex-analytic cycles in V which are invariant under conjugation (this set includes the complexifications of all real-analytic cycles in V^0). This mapping α is defined for irreducible subvariety X by setting $\alpha(X) = X^0 = X \cap V^0$ in case

X^0 has the real dimension r, and setting $\alpha(X) = 0$ otherwise; this definition is then extended to complex-analytic cycles in V by linearity and reduction mod 2. The composite mapping $\psi \circ \alpha$ of complex-analytic cycles in V into $H_*(V^0, \mathbf{Z}_2)$ is then the mapping ρ mentioned in Section 1. For our present purpose it is significant to observe that a complex-analytic variety V can also be considered as a real-analytic variety, and the complex-analytic cycles in V, taken over $\mathbf{Z}_2$, are also real analytic cycles; and that the intersection theory for real-analytic cycles over $\mathbf{Z}_2$ in V as a real-analytic variety, when restricted to the complex-analytic cycles in V over $\mathbf{Z}_2$, coincides with the intersection theory for complex-analytic cycles in V as a complex-analytic variety, taken mod 2.

THEOREM 1. *Let V be a complex-analytic variety of complex dimension n which is the complexification of a real-analytic variety V^0, and let X be a complex-analytic cycle in V such that X and $\bar{X}$ intersect properly in V at every component containing a real point; then X and V^0 intersect properly in V as real-analytic cycles over $\mathbf{Z}_2$, and we have the relation $\alpha(X \cdot \bar{X}) = X \cdot V^0$.*

Proof. Since the properties asserted here are linear over $\mathbf{Z}_2$, it is sufficient to restrict ourselves to the case where X is a complex-analytic subvariety in V. Let n and r denote the complex dimensions of V and X respectively. Let p be a real point in $X \cap \bar{X}$, and let $z_j = x_j + iy_j$, $j = 1, \cdots, n$, be a system of local coordinates of V at p such that $x_1, \cdots, x_n$ form a system of local coordinates of V^0 at p. Let X be defined in the neighborhood of p by the equations $f_k(z) = g_k(z) + ih_k(z) = 0$, $k = 1, \cdots, s$, where f_k are convergent power series with complex coefficients and g_k and h_k are convergent power series with real coefficients; then $\bar{X}$ is defined by the equations $\bar{f}_k(z) = g_k(z) - ih_k(z) = 0$. Then the analytic set $X \cap \bar{X}$ is determined in the neighborhood of p by the equations $g_k(z) = h_k(z) = 0$, and our assumption asserts that every component of the solutions of these equations has the complex dimension $2r - n$. Now, let V^* be a complexification of V considered as a real-analytic variety of real dimension $2n$, and let $x_1^*, y_1^*, \cdots, x_n^*, y_n^*$ be a system of local coordinates of V^* at p such that $x_1, y_1, \cdots, x_n, y_n$ constitute its real part. If we set $g_k(z) = g_k'(x, y) + ig_k''(x, y)$ and $h_k(z) = h_k'(x, y) + ih_k''(x, y)$, then the complexification X^* of X in V^* is determined locally at p by the equations $g_k'(x^*, y^*) - h_k''(x^*, y^*) = g_k''(x^*, y^*) + h_k'(x^*, y^*) = 0, k = 1, \cdots, s$; and it is clear that the complexification V^{0*} of V^0 in V^* is determined locally at p by the equations $y_1^* = \cdots, y_n^* = 0$.

If we observe that $g_k''(x, 0) = h''(x, 0) = 0$ and hence also $g_k''(x^*, 0) = h_k''(x^*, 0) = 0$, then the analytic set $X^* \cap V^{0*}$ is determined locally at p

by the equations $g_k'(x^*,0) - h_k'(x^*,0) = 0$, $k = 1,\cdots,s$. On the other hand, since $g_k(z)$ and $h_k(z)$ have real coefficients, they remain unchanged if we first restrict the variables z to their real parts x and then replace these real variables x again by the complex variables z, i.e. $g_k(z) = g_k'(z,0)$ and $h_k(z) = h_k'(z,0)$. Replacing the variables z by x^* it follows that every component of the solutions of the equations $g_k'(x^*,0) - h_k'(x^*,0) = 0$, $k = 1,\cdots,s$, in the neighborhood of p has the complex dimension $2r - n$. This shows that X^* and V^{0*} intersect properly in V^* at every component containing a real point, and by definition, this means that X and V^0 intersect properly in V.

To prove the relation $\alpha(X\cdot\bar{X}) = X\cdot V^0$, we shall show that both sides have the same multiplicity at each component. We need only consider those components of $X \cap V^0$ having a real dimension $2r - n$, for any component of a lower dimension will have complexifications in V and in V^* which are not proper components in $X \cap \bar{X}$ and $X^* \cap V^{0*}$ respectively and hence will have by definition the multiplicity zero in both $\alpha(X,\bar{X})$ and $X\cdot V^0$. Let C^0 then be such a component, and let C and C^{0*} be its complexifications in V and V^* respectively, so that they are proper components in $X \cap \bar{X}$ and $X^* \cap V^{0*}$ respectively; let p be a point in C^0 which is simple in C (hence also in C^0) and does not lie in any other component of $X \cap V^0$. If L is a local linear subvariety at p of complex dimension $2n - 2r$, defined over $\boldsymbol{R}$ (i.e. by setting to zero $2r - n$ linear forms of the local coordinates z with real coefficients), which is transversal to C at p, then the intersection cycle $L\cdot X$ is defined locally in a neighborhood of p, and we have the relations $i((L\cdot X)\cdot(\overline{L\cdot X}),p)_L = i(X\cdot\bar{X},C)$ and $i((L\cdot X)\cdot L^0,p)_L = i(X\cdot V^0,C^0)$. Since our problem is essentially local (i.e. we only need to consider X as defined in a neighborhood of p), this shows that we can restrict ourselves to the case where $2r - n = 0$ and hence $C = C^0 = p$. According to Lemma 1, the multiplicity $i(X\cdot\bar{X},p)$ of $X\cdot\bar{X}$ at p is equal to the number of solutions in a suitable neighborhood N of the origin of the equations $f_k(z + 2ie) = \bar{f}_k(z) = 0$, $k = 1,\cdots,s$, for almost all sufficiently small real n-vectors e. If we perform a local coordinate translation by replacing z by $z + ie$, then the equations will have the form $f_k(z + ie) = \bar{f}_k(z - ie) = 0$, $k = 1,\cdots,s$, but the number of solutions in N will remain the same provided e is taken sufficiently small. If we set $F_k(z) = f_k(z + ie)$ then our equations will have the form $F_k(z) = \bar{F}_k(z) = 0$, $k = 1,\cdots,s$; hence, if we set $F_k(z) = G_k(z) + iH_k(z)$, where G_k and H_k are power series with real coefficients, then our equations can be written in the form $G_k(z) = H_k(z) = 0$, $k = 1,\cdots,s$. On the other hand, again by Lemma 1, the multiplicity $i(X^*\cdot V^{0*},p)$ is equal to the number of

solutions in N (replacing N by a smaller neighborhood if necessary) of the equations $g_k'(x^*, e) - h_k''(x^*, e) = g_k''(x^*, e) + h_k'(x^*, e) = 0$, $k = 1, \cdots, s$, for almost all sufficiently small real n-vectors e. If we set

$$G_k(z) = G_k'(x, y) + iG_k''(x, y) \text{ and } H_k(z) = H_k'(x, y) + iH_k''(x, y),$$

then we have the relations $G_k'(x, y) - H_k''(x, y) = g_k'(x, y + e) - h_k''(x, y + e)$, and $G_k''(x, y) + H_k'(x, y) = g_k''(x, y + e) + h_k'(x, y + e)$, since both sides are the real and imaginary parts respectively of the function $F_k(z) = f_k(z + ie)$. It follows that $G_k'(x, 0) = g_k'(x, e) - h_k''(x, e)$ and $H_k'(x, 0) = g_k''(x, e) + h_k'(x, e)$, and hence $G_k'(x^*, 0) = g_k'(x^*, e) - h_k''(x^*, e)$ and $H_k'(x^*, 0) = g_k''(x^*, e) + h_k'(x^*, e)$, so that $i(X^* \cdot V^{0*}, p)$ is equal to the number of solutions in N of the equations $G_k'(x^*, 0) = H_k'(x^*, 0) = 0$, $k = 1, \cdots, s$. Now, since both G_k and H_k are real, a similar argument as before shows that $G_k(z) = G_k'(z, 0)$ and $H_k(z) = H_k'(z, 0)$, so that $i(X^* \cdot V^{0*}, p)$ is equal to the number of solutions in N of the equations $G_k(z) = H_k(z) = 0$, $k = 1, \cdots, s$. This shows that $i(X \cdot \bar{X}, p) = i(X^* \cdot V^{0*}, p)$; taking both sides mod 2, we conclude that $i(X \cdot V^0, p)$ is equal to the multiplicity of p in $\alpha(X \cdot \bar{X})$. This concludes the proof of Theorem 1.

Remark. The equation $i(X \cdot \bar{X}, p) = i(X^*, V^{0*}, p)$ shows that X and $\bar{X}$ are transversal at p in V if and only if X^* and V^{0*} are transversal at p in V^*, which is so if and only if X and V^0 are transversal at p in V. This can be easily verified by direct calculation of the equations of the tangent spaces of X, $\bar{X}$, and V^0 at p.

4. From now on we shall assume that V is algebraic and quasi-projective, and that the cycles involved are also algebraic. The reason for making this restriction is that the algebraic cycles in a quasi-projective algebraic variety can be "moved around" up to equivalence so that any two or finite number of them can be made to intersect properly (in fact even transversally, but we shall not need this). We shall state the properties we need here in a lemma as follows.

LEMMA 2. *Let V be a quasi-projective complex-algebraic variety, defined over $\mathbf{R}$, which is a complexification of its real trace V^0, and let X be a complex-algebraic cycle in V which is defined over $\mathbf{R}$; then*

(a) *there exists a complex-algebraic cycle X', defined over $\mathbf{R}$ and rationally equivalent to X over $\mathbf{R}$, such that X' intersect properly in V any given cycle Y which is defined over $\mathbf{R}$ (or over $\mathbf{C}$), and*

(b) *there exists a complex-algebraic cycle X'', rationally equivalent to X, such that X'' and $\overline{X''}$ intersect properly in V.*

Proof. We shall apply a result in [3]; for this purpose we introduce a universal domain $\boldsymbol{K}$ which contains $\boldsymbol{C}$ as an "admissible" field, and denote by $V_{\boldsymbol{K}}$, $X_{\boldsymbol{K}}$, $Y_{\boldsymbol{K}}$, the extensions of V, X, Y, respectively over $\boldsymbol{K}$, so that $V_{\boldsymbol{K}}$, $X_{\boldsymbol{K}}$, $Y_{\boldsymbol{K}}$, are algebraic variety and cycles respectively in the sense of [3] and are defined over $\boldsymbol{R}$, and V, X, Y are obtained from $V_{\boldsymbol{K}}$, $X_{\boldsymbol{K}}$, $Y_{\boldsymbol{K}}$, respectively by restriction to points which are rational over $\boldsymbol{C}$. We now apply Lemma 2, Corollary 4, in [3] with $\boldsymbol{R}$ as the ground field, and obtain a cycle $X(u)$ in V, defined over a purely transcendental extension $\boldsymbol{R}(u) = \boldsymbol{R}(u_1, \cdots, u_m)$ of $\boldsymbol{R}$ and rationally equivalent to $X_{\boldsymbol{K}}$ over $\boldsymbol{R}(u)$, such that $X(u)$ intersects $Y_{\boldsymbol{K}}$ properly in $V_{\boldsymbol{K}}$ and specializes to $X_{\boldsymbol{K}}$ over the specialization $u \to 0$ over $\boldsymbol{R}$. Let $f_i(u)$ be a system of elements in $\boldsymbol{R}[u]$ which represent the associated coordinates of the cycle $X(u)$, and let μ be a real m-vector which is not a common zero of $f_i(u)$; then there is a unique specialization $X(\mu)$ of $X(u)$ over the specialization $u \to \mu$ over $\boldsymbol{R}$, and $X(\mu)$ is defined over $\boldsymbol{R}$ and is rationally equivalent to $X_{\boldsymbol{K}}$ over $\boldsymbol{R}$. Now, the condition for two algebraic cycles in $V_{\boldsymbol{K}}$ to intersect improperly can be expressed by the vanishing of a finite number of polynomials in the associated coordinates of the cycles with coefficients in $\boldsymbol{R}$. There exists therefore a set of polynomials $g_j(u)$ in $\boldsymbol{R}[u]$ such that $X(\mu)$ intersects $Y_{\boldsymbol{K}}$ properly in $V_{\boldsymbol{K}}$ if and only if μ is not a common zero of $g_j(u)$. If we choose the real m-vector μ so that it is not a common zero of the set $f_i(u)g_j(u)$, and denote by X' the restriction of $X(\mu)$ to rational points over $\boldsymbol{C}$, then X' is a cycle in V satisfying (a). Next, let μ be a complex m-vector which is not a common zero of $f_i(u)$; then there exist unique specializations $X(\mu)$ and $X(\bar{\mu})$ of $X(u)$ over the specializations $u \to \mu$ and $u \to \bar{\mu}$ respectively over $\boldsymbol{R}$, and both $X(\mu)$ and $X(\bar{\mu})$ are defined over $\boldsymbol{C}$ and are rationally equivalent to $X_{\boldsymbol{K}}$ over $\boldsymbol{C}$; and it is clear that $X(\bar{\mu})$ is the conjugate of $X(\mu)$. If we now take a system of variables $v = (v_1, \cdots, v_m)$ over $\boldsymbol{R}$ in $\boldsymbol{K}$ which are independent with respect to u over $\boldsymbol{R}$, then the cycle $X(v)$ has the same properties as $X(u)$, except that the variables u are now replaced by v; furthermore, again by [3], Lemma 2, Corollary 4, the cycles $X(u)$ and $X(v)$ intersect properly in $V_{\boldsymbol{K}}$. It follows then, just as before, that there exists a set of polynomials $h_k(u, v)$ in $\boldsymbol{R}[u, v]$ such that $X(\mu)$ and $X(\nu)$ intersect properly in $V_{\boldsymbol{K}}$ if and only if the pair μ, ν is not a common zero of $h_k(u, v)$. If we choose the complex m-vector μ so that the pair μ, $\bar{\mu}$ is not a common zero of the set $f_i(u)h_k(u, v)$ (or equivalently of the set $f_i(u)f_j(v)h_k(u, v)$), and denote by X'' the restriction of $X(\mu)$ to rational points over $\boldsymbol{C}$, then X'' is a cycle in V satisfying (b). This concludes the proof of Lemma 2.

We are now in a position to prove the result announced at the beginning of this paper.

THEOREM 2. *Let V be a quasi-projective complex-algebraic variety, defined over $\mathbf{R}$, which is a complexification of its real trace V^0, and let X be a complex-algebraic cycle in V which is defined over $\mathbf{R}$; then $\phi(X)\cdot\phi(V^0)$ and $\rho(X)^2$, both as elements in $H_*(V^0,\mathbf{Z}_2)$, are numerically equivalent with respect to real-algebraic homology classes in $H_*(V^0,\mathbf{Z}_2)$.*

Proof. By Lemma 2(a), there exists a complex-algebraic cycle X' in V, defined over $\mathbf{R}$ and rationally equivalent to X over $\mathbf{R}$, such that X and X' intersect properly. It follows then from [1], Theorem 5.3, that

$$\rho(X\cdot X')=\rho(X)\cdot\rho(X')=\rho(X)^2.$$

On the other hand, by Lemma 2(b), there exists a complex-algebraic cycle X'' in V, rationally equivalent to X, such that X'' and $\overline{X''}$ intersect properly in V. It follows then from Theorem 1 that $X''\cdot V^0=\alpha(X''\cdot\overline{X''})$ and hence $\phi(X)\cdot\phi(V^0)=\phi(X'')\cdot\phi(V^0)=\rho(X''\cdot\overline{X''})$. It remains only to show that $\rho(X\cdot X')$ and $\rho(X''\cdot\overline{X''})$ are numerically equivalent with respect to real-algebraic homology classes in $H_*(V^0,\mathbf{Z}_2)$. Let Y^0 be a real-algebraic cycle of the complementary real dimension to $\alpha(X\cdot X')$ in V^0 (i.e., if n and r are the real dimension of V^0 and $\alpha(X)$ respectively, then the real dimension of Y^0 should be $n-(2r-n)=2n-2r$), and let Y be its complexification in V; in view of Lemma 2(a), we can assume that Y intersect both $X\cdot X'$ and $X''\cdot\overline{X''}$ properly in V. Since $X\cdot X'$ and $X''\cdot\overline{X''}$ are rationally equivalent, they must have the same intersection number with Y; it follows that $\alpha(X\cdot X')$ and $\alpha(X''\cdot\overline{X''})$ must have the same intersection number with $Y^0=\alpha(Y)$, and hence $\rho(X\cdot X')$ and $\rho(X''\cdot\overline{X''})$ must have the same intersection number with $\psi(Y^0)$. This proves our theorem.

We conclude this paper by mentioning one further application of our result. Assume that $H_n(V,\mathbf{Z}_2)$ has a numerical equivalence base $B_1,\cdots,B_t$ consisting of complex-algebraic cycles defined over $\mathbf{R}$, and let $M=(m_{ij})$ be its intersection matrix; then the cycle V^0 is numerically equivalent over $\mathbf{Z}_2$ to a cycle of the form $\sum_{i=1}^{t} n_iB_i$. Now, Theorem 2 provides us with the numerical relations $\phi(V^0)\cdot\phi(B_j)=\rho(B_j)^2$, or $\sum_{i=1}^{t} n_im_{ij}=m_{jj}$, $j=1,\cdots,t$, from which we can easily calculate the n_i and hence the self intersection number $\phi(V^0)^2$ as a cycle over $\mathbf{Z}_2$ in V. For example, in case V is the complex Grassmannian, it is well-known that $H_n(V,\mathbf{Z}_2)$ has a base consisting of the Schubert varieties

of complex dimension $n/2$ (none in case n is odd), each of which is the dual of some other or itself, and a simple calculation shows that the number $\phi(V^0)^2$ is equal mod 2 to the total number of Schubert varieties of dimension n, or equivalently the number of those which are self-dual.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. Borel and A. Haefliger, "La classe d'homologie fondamentale d'un espace analytique," *Bulletin de la Société Mathématique de France*, vol. 89 (1961), pp. 461-513.

[2] A. Borel and F. Hirzebruch, "Characteristic classes and homogeneous spaces, II," *American Journal of Mathematics*, vol. 81 (1959), pp. 315-382, see Appendix II.

[3] W. L. Chow, "On equivalence classes of cycles in an algebraic variety," *Annals of Mathematics*, vol. 64 (1956), pp. 450-479.

[4] H. Cartan, "Variétés analytiques-réelles et variétés analytiques complexes," *Bulletin de la Société Mathématique de France*, vol. 85 (1957), pp. 77-100.

[5] H. Whitney and F. Bruhat, "Quelques propriétés fondamentales des ensembles analytiques-réels," *Commentarii Mathematici Helvetici*, vol. 33 (1959), pp. 132-160.

AUTOMORPHISMS OF COMPACT RIEMANN SURFACES.*[1]

By JOSEPH LEWITTES.

Introduction. An automorphism of a Riemann surface is a conformal homeomorphism of the surface onto itself. In this paper we investigate the properties of automorphisms of compact Riemann surfaces. For such surfaces with genus greater than one the group of automorphisms is a finite group. The result is due to Hurwitz [4] and is proved in the first section where we collect for reference most of the terms and theorems useful for our purposes. In fact the two important papers of Hurwitz [3], [4] have been the impetus for many of the topics that we consider. In the second section we discuss the representation of the group of automorphisms as linear transformations of the spaces of differentials and determine the dimension of the space of analytic differentials invariant under a group of automorphisms. In the third section we examine the diagonal form of the matrix of the representation in the case of a cyclic group of prime order. In the fourth section a new method enables one to obtain explicitly the matrix of the representation and leads to a striking result concerning the number of fixed points of an automorphism. Namely if a point which is not a Weierstrass point is fixed, then there are only two, three or four fixed points. We then apply the general results to obtain specific information about the automorphisms of two particular types of surfaces, hyperelliptic and those having $(g-1)g(g+1)$ Weierstrass points, where g is the genus. A related result is that if S/H is the quotient space of a surface S under a group H of automorphisms of prime order N, then the genus of S/H is simply $[\frac{g}{N}]$, the greatest integer in $\frac{g}{N}$, which is independent of the Riemann-Hurwitz relation for it does not involve the number of fixed points.

I now wish to take the opportunity to express my gratitude and indebtedness to Dr. H. E. Rauch who directed my research and was the source of many fruitful ideas.

* Received December 3, 1962.

[1] This paper is based on the author's doctoral thesis submitted to Yeshiva University, 1962. During his pre-doctoral studies the author held a U. S. Steel Foundation Fellowship. Work on this paper was partially supported by NSF G 18929.

1. In this preliminary section we present the notation and a concise review of some of the theorems and concepts to be used later. As general references for the theory of Riemann surfaces one may consult [1], [8]. We concern ourselves with compact Riemann surfaces only. Let S be a Riemann surface of genus g; we denote by D_m the complex vector space of meromorphic differentials of type m on S, $m \geqq 0$, and by $A_m \subset D_m$ the subspace consisting of analytic differentials. For example A_1 is the space of Abelian differentials of the first kind and D_0 is the space of meromorphic functions on S, which we simply call functions. It is known that A_1 has dimension g and any $\theta \in A_1$, not identically zero, has $2g-2$ zeros while for $m > 1$, $\dim A_m = (2m-1)(g-1)$ and any $\theta \in A_m$, not identically zero, has $m(2g-2)$ zeros.

If $\pi\colon S_1 \to S_2$ is an analytic branched covering, S_i having genus g_i $(i=1,2)$, then the Riemann-Hurwitz relation states that

$$2g_1 - 2 = N(2g_2 - 2) + B$$

where N is the number of 'sheets' in S_1 over S_2 and B is the sum of the orders of the branch points on S_1.

(θ) denotes the divisor of zeros and poles of $\theta \in D_m$. Concerning divisors on S the Riemann-Roch theorem states $r(b^{-1}) = d(b) + i(b) + 1 - g$, where $r(b^{-1})$ is the dimension of the space of functions which are multiples of b^{-1}, $d(b)$ the degree of b, and $i(b)$ the dimension of the subspace of D_1 consisting of multiples of b. As a consequence of this theorem one has the Weierstrass gap theorem which may be formulated as follows. To each point $P \in S$, $g > 0$, there may be assigned a sequence of integers $\alpha(P) = (\alpha_1(P), \cdots, \alpha_g(P))$ with $0 < \alpha_1$, $\alpha_i < \alpha_{i+1}$, $\alpha_g = 2g$, such that for each α_i there exists a function on S with exactly α_i poles at P and no others, and that if a function on S has poles only at P, these being $n \leqq 2g$ in number, then $n = \alpha_i$ for some i. Such a function will be called a function at $\alpha_i(P)$. The sequence complementary to $\alpha(P)$ in the sequence of integers from 1 to $2g$ is called the *gap* sequence, denoted by $\gamma(P) = (\gamma_1(P), \cdots, \gamma_g(P))$ where $0 < \gamma_1$, $\gamma_i < \gamma_{i+1}$, $\gamma_g < 2g$.

$P \in S$ is called a Weierstrass point if $\alpha_1(P) \leqq g$. Hurwitz [4] has shown that if S has $g \geqq 2$ then there do exist Weierstrass points, their number ω is finite, bounded by $2g+2 \leqq \omega \leqq (g-1)g(g+1)$. Denoting by W the set of Weierstrass points we see that $P \notin W$ if and only if $\gamma(P) = (1, 2, \cdots, g)$. Also, one may prove that for any $P \in S$ there exists a basis $\theta_1, \cdots, \theta_g$ of A_1 such that θ_i has a zero of order $\gamma_i - 1$ at P. We call this a basis of A_1 at P.

S is called hyperelliptic if $g \geqq 2$ and there is a function on S with only two poles. This is easily seen to be equivalent to saying that for some $P \in S$ $\alpha_1(P) = 2$. It can then be shown that S is hyperelliptic if and only if

$\omega = 2g + 2$ and that at each $P \in W$, $\alpha(P) = (2, 4, \cdots, 2g)$ and $\gamma(P) = (1, 3, \cdots, 2g - 1)$.

An automorphism of S is a conformal homeomorphism of S onto itself. Let $H(S)$ denote the group of automorphisms of S. If $g = 0$ or 1 then it is clear that $H(S)$ is not finite. On the other hand, if $g \geqq 2$ Hurwitz [4] proved that $H(S)$ is a finite group. We sketch the proof of this because it motivates some of the ideas which appear later. First, note that $\alpha(P)$, $\gamma(P)$ are invariants under $H(S)$, for if f is a function at $\alpha_i(P)$ then fh^{-1} is a function at $\alpha_i(h(P))$; $h \in H(S)$. In particular, h takes W into itself, giving us a representation of $H(S)$ as a permutation group by means of the homomorphism $L: H(S) \to \Sigma_\omega$, Σ_ω being the symmetric group on ω letters, defined by $L(h) = h \mid W$. Σ_ω being a finite group we need only show that the kernel of L is finite to conclude $H(S)$ is finite. To this end we must examine the set of fixed points of $h \in H(S)$, points $P = h(P)$. If $h \neq I$, the identity, the set of fixed points of h being closed and W a finite set, we may find $P \neq h(P)$, $P \notin W$. Let f be a function at $\alpha_1(P)$ so that it has $g + 1$ poles at P and no others. Then $f_1 = f - fh$ has $g + 1$ poles at P and $g + 1$ poles at $h(P) \neq P$ and is otherwise finite; hence f_1 is not a constant. Clearly every fixed point of h is a zero of f_1, which, having $2g + 2$ poles, has $2g + 2$ zeros. Thus h cannot have more than $2g + 2$ fixed points.

Returning to the representation $L: H(S) \to \Sigma_\omega$, if S is not hyperelliptic then $\omega > 2g + 2$ and our above result on the number of fixed points not exceeding $2g + 2$ implies in fact that L is a faithful representation or that L is an isomorphism (into). If S is hyperelliptic, then it is well known that S admits an automorphism J of order 2 whose fixed points are precisely the $\omega = 2g + 2$ points of W. J is simply the interchange of the two sheets of S mapped over the sphere by the function with two poles on S which characterizes the hyperelliptic surface. J is uniquely determined by its order and number of fixed points. In fact if S is any surface admitting an automorphism h of order 2 with $2g + 2$ fixed points, then S is hyperelliptic and $h = J$. This is an easy consequence of the Riemann-Hurwitz relation. In any case, we conclude that when S is hyperelliptic the kernel of L is (I, J)—showing this is a normal subgroup—and we again obtain that $H(S)$ is finite.

Finally, let us observe that the bound $2g + 2$ for the number of fixed points of h can be sharpened. For if $h \neq I$, and if S hyperelliptic, $h \neq J$, then our above results tell us that there is a point $P \neq h(P)$, $P \in W$. Among all such P choose one whose $\alpha_1(P)$ is minimal. Let f be a function at $\alpha_1(P)$ and consider $f_1 = f - fh$. It is now immediate that h has at most $2\alpha_1(P)$ fixed points. Of course this bound depends on h, but since $\alpha_1(P) \leqq g$ we

obtain that for any $h \neq I, J$ there are at most $2g$ fixed points while if S is hyperelliptic, since $\alpha_1(P) = 2$, there are at most 4 fixed points. In [4], Hurwitz only points out the latter result for a hyperelliptic S but not the more general result.

2. The representation of $H(S)$ as permutations of W has yielded quite a bit of information. One is then led naturally to consider the representation of $H(S)$ as linear transformations of the spaces D_m and A_m. Explicitly this is done as follows. Let $\theta \in D_m$ have the representation at $P \in S$ in terms of a local parameter z at P as $\theta = (a_k z^k + \cdots) dz^m$. For $h \in H(S)$, let $P_1 = h^{-1}(P)$ and say θ at P_1 in terms of a local parameter u at P_1 is $\theta = (b_j u^j + \cdots) du^m$. Since h^{-1} is conformal, the mapping h^{-1} is given locally at $P_1 = h^{-1}(P)$ as $u = F(z)$, $F'(0) \neq 0$. We now define $h(\theta)$ to be the differential which at P in z is given as $h(\theta) = (b_j(F(z))^j + \cdots)(F'(z))^m dz^m$. Concisely, we may write $h(\theta) = \theta h^{-1}$ and it is immediate that $h(\theta)$ has behavior with respect to zeros and poles at $h(P)$ as does θ at P. Clearly $\theta \rightarrow h(\theta)$ is a linear transformation of D_m onto itself which takes A_m onto A_m. For clarification let us denote by R_m the representation $h \rightarrow R_m(h)$ where $R_m(h)$ is the linear transformation determined by h of A_m onto A_m.

THEOREM 1. *R_m is faithful if $m \geqq 1$ except when $g = 2$ and m is even in which case the kernel of R_m is (I, J).*

Proof. Consider first R_1 and assume $h \neq I$. If S is hyperelliptic and $h = J$ then it is well known that $J(\theta) = -\theta$ for all $\theta \in A_1$ (this can be seen by writing down the differentials explicitly but we shall give an independent proof of this in a later section) so that $R_1(J)$ is not the identity on A_1. If now $h \neq J$ and S is arbitrary, there is a $P \in W$ such that $h(P) \neq P$ and a $\theta \in A_1$ with at least g zeros at P since there is a gap at P greater than g. Then $h(\theta)$ has at least g zeros at $h(P)$ so that $h(\theta) = \theta$ implies θ has at least $2g$ zeros which is absurd. Thus $h(\theta) \neq \theta$ and so $R_1(h)$ is not the identity if $h \neq I$. If $m > 1$, using the same notation and argument we conclude that $\theta^m \in A_m$ cannot satisfy $h(\theta^m) = \theta^m$ so that $R_m(h)$ is the identity on A_m only if $h = I$. If $m > 1$ and $h = J$ then for $g > 2$ one observes that there is always a $\theta \in A_m$ such that $J(\theta) = -\theta$ while if $g = 2$ and $m > 1$ then $J(\theta) = (-1)^m \theta$ for all $\theta \in A_m$. This completes the proof.

COROLLARY 1. *$H(S)$ is represented faithfully on $H_1(S)$, the first homology group.*

Proof. Let a_j, b_j, $1 \leqq j \leqq g$ be $2g$ cycles of a canonical dissection of S which generate the first (simplicial) homology group. Again $a_j \rightarrow h(a_j)$,

$b_j \to h(b_j)$—where $h(a_j)$, $h(b_j)$ is simply the cycle consisting of the point set image of a_j, b_j resp. under h—is a representation of $H(S)$. Let $\theta_1, \cdots, \theta_g \in A_1$ be a normalized basis so that $\int_{a_j} \theta = \delta_{ij}$ (Kronecker delta). Now $\int_{h(a_j)} h(\theta_i) = \int_{a_j} \theta_i$—since $h(\theta_i) = \theta_i h^{-1}$—so that if h is the identity on $H_1(S)$, $h(a_j)$ is homologous to a_j and so $\int_{a_j} h(\theta_i) = \delta_{ij}$ and so $h(\theta_i) = \theta_i$ since the normalized basis is uniquely determined. But $h(\theta_i) = \theta_i$ for $1 \leqq i \leqq g$ implies $h = I$ by the theorem.

COROLLARY 2. *If $h_1, h_2, \in H(S)$, $h_1 \neq h_2$ then h_1 is not homotopic (topologically) to h_2.*

Proof. For if homotopic, then $h_1 h_2^{-1}$ is the identity on $H_1(S)$ and then by the previous corollary $h_1 = h_2$.

Now let $H \subset H(S)$ be a subgroup. Relating $P_1, P_2 \in S$ if $P_2 = h(P_1)$ for some $h \in H$ is an equivalence relation and it is not difficult to see that the quotient space $\bar{S} = S/H$ may be given the obvious conformal structure with respect to which $\bar{S}$ is a Riemann surface and $\pi: S \to \bar{S}$ is a branched analytic covering whose branch points are at those points of S which are fixed under some $h \in H$. In fact π defines S as an N sheeted covering of $\bar{S}$ where N is the order of H and a point $P \in S$ has branch order $N(P) - 1$ where $N(P)$ is the order of the subgroup of H leaving P fixed, which we call the stability group of P. If $\pi(P_1) = \pi(P_2) = \bar{P} \in \bar{S}$ then P_1, P_2 have conjugate stability groups and so $N(P_1) = N(P_2)$. In this way we may define unambiguously $N(\bar{P})$ as $N(P)$ where P is any point such that $\pi(P) = \bar{P}$. With this notation we observe that $\bar{P} = \pi(P)$ may be written in terms of suitable local parameters $\bar{z}$, z at $\bar{P}$, P resp. as $\bar{z} = z^{N(P)} = z^{N(\bar{P})}$. We denote the genus of $\bar{S}$ by $\bar{g}$ and the spaces of differentials on $\bar{S}$ as $\bar{D}_m$ and $\bar{A}_m$. $\theta \in D_m$ is called H invariant if $h(\theta) = \theta$ for all $h \in H$. The set of H invariant differentials form a subspace $D_m{}^H \subset D_m$ and similarly we have a subspace $A_m{}^H = D_m{}^H \cap A_m$. If $\bar{\theta} \in \bar{D}_m$ it may be 'lifted' to a differential $\theta = \pi^{-1}(\bar{\theta}) \in D_m{}^H$ as follows. Let $\bar{P} = \pi(P)$ be locally $\bar{z} = z^{N(P)}$ and say $\bar{\theta}$ at $\bar{P}$ in $\bar{z}$ is $\bar{\theta} = (a_k \bar{z}^k + \cdots) d\bar{z}^m$. Then define θ at P in z as $\theta = (a_k (z^{N(P)})^k + \cdots)(\frac{d\bar{z}}{dz}) dz^m$. By the nature of this construction $h(\theta) = \theta$ for all $h \in H$ and so $\theta \in D_m{}^H$ and $\pi^{-1}: \bar{D}_m \to D_m{}^H$ is in fact an isomorphism (into). Observe that if $\bar{\theta}$ at P has a zero or pole of order $k (\geqq 0)$ then $\pi^{-1}(\bar{\theta})$ at P has correspondingly a zero of order $(k+m)N(P) - m$ or pole of order $(k-m)N(P) + m$.

To show that π^{-1} is in fact onto $D_m{}^H$ we 'lower' differentials from $D_m{}^H$ to $\tilde{D}_m$. For this we note that the stability group of any $P \in S$ is cyclic. Indeed, one can find a suitable parameter z at P such that any h leaving P fixed is locally $z \to \mu(h)z$ where $|\mu(h)| = 1$ is a primitive $N(h)$-th root of unity, $N(h)$ being the order of h. The map $h \to \mu(h)$ is then clearly an isomorphism of the stability group of P onto a finite abelian group of roots of unity. But a group of this latter type is necessarily cyclic. Now if $\theta \in D_m{}^H$, $P \in S$, h of order $N(h) = N(P) > 1$ generates the cyclic stability group of P, z a local parameter at P such that h^{-1} is locally $z \to \mu z$, μ primitive $N(P)$-th root of unity let θ at P in z be given by $\theta = (a_n z^n + \cdots) dz^m$. The hypothesis $\theta \in D_m{}^H$ means $h(\theta) = \theta$ so that we must have

$$(a_n z^n + \cdots) dz^m = (a_n (\mu z)^n + \cdots) \mu^m dz^m$$

Thus the only non zero coefficients a_k which can appear are when $k + m \equiv 0 \pmod{N(P)}$ and in fact θ appears as

$$(1) \qquad (a_{nN(P)-m} z^{nN(P)-m} + \cdots + a_{(n+k)N(P)-m} z^{(n+k)N(P)-m} + \cdots) dz^m.$$

Now if $P_1 \in S$ and $\pi(P_1) = \pi(P)$, because θ is H invariant we may find a parameter z at P_1 in which θ again has the expansion (1). Thus if $\tilde{P} = \pi(P)$ we may unambiguously define a differential $\tilde{\theta} = \pi(\theta)$ which in a local parameter $\tilde{z} = z^{N(P)}$ at $\tilde{P}$ is

$$(2) \qquad \begin{aligned} \tilde{\theta} &= (a_{nN(P)-m} z^{nN(P)-m} + \cdots) \frac{1}{\left(\dfrac{d\tilde{z}}{dz}\right)^m} d\tilde{z}^m \\ &= \frac{1}{N(P)m} (a_{nN(P)-m} \tilde{z}^{(n-m)} = \cdots) d\tilde{z}^m \end{aligned}$$

Then $\theta \to \pi(\theta)$ is a linear transformation of $D_m{}^H$ into $\tilde{D}_m$ with inverse π^{-1}. We have therefore established:

THEOREM 2. $\pi \colon D_m{}^H \cong \tilde{D}_m$.

This isomorphism can be regarded as an identification of the two spaces for it does not involve any choice of basis, i.e. it is canonical.

How do the spaces $A_m{}^H$, $\tilde{A}_m$ behave with respect to this isomorphism? If $\tilde{\theta} \in \tilde{A}_m$ then by our previous observations it is clear that $\pi^{-1}(\tilde{\theta}) \in A_m{}^H$. However, if $\theta \in A_m{}^H$ is (1) above, then $nN(P) - m \geqq 0$ but $\tilde{\theta} = \pi(\theta)$ starts out like $\tilde{z}^{n-m}$ which may be a pole. Conversely we see that a $\tilde{\theta} \in \tilde{D}_m$ having poles may lift to $\pi^{-1}(\tilde{\theta}) \in A_m{}^H$. However in either case the poles can occur only at the finite set $\tilde{P}_1, \cdots, \tilde{P}_t$, of branch points on $\tilde{S}$, i.e. those $\tilde{P}$ with $N(\tilde{P}) > 1$. Thus if H has no fixed points, $\pi \colon S \to \tilde{S}$ is smooth and in this case π:

$A_m{}^H \cong \tilde{A}_m$. Assuming now there are $t > 0$ branch points on $\tilde{S}$, define, for any integer m and $\tilde{P} = \pi(P)$, the integer $(P, m) = (\tilde{P}, m)$ as the least integer such that

$$(P, m)\, N(P) - m \geqq 0.$$

Then $(P, m) = m$ whenever $N(P) = N(\tilde{P}) = 1$ while at a branch point, since $N(P) = N(\tilde{P}) \geqq 2, 1 \leqq (P, m) \leqq m/2 + 1$. Hence from (1) and (2) above, we see that if $\theta \in A_m{}^H$ then $\pi(\theta)$ has at most a pole of order $m - (\tilde{P}_i, m)$ at $\tilde{P}_i \in \tilde{S}$, $1 \leqq i \leqq t$ and is otherwise regular. Conversely any such $\tilde{\theta}$ lifts to $\pi^{-1}(\tilde{\theta}) \in A_m{}^H$. When $m = 1$, $(P, 1) = (\tilde{P}, 1) = 1$ for all points so that in this case no poles can appear and $\pi\colon A_1{}^H \cong \tilde{A}_1$. Assume now $m \geqq 2$. We have already characterized the isomorphic image of $A_m{}^H$ under π and so we may compute the dimension of $A_m{}^H$ as the dimension of the space of differentials $\tilde{\theta} \in \tilde{D}_m$ which have at most poles of order $m - (\tilde{P}_i, m)$ at $\tilde{P}_i$, $1 \leqq i \leqq t$ and no others. This is a standard application of the Riemann-Roch theorem as follows. Assume $\tilde{\theta} \in \tilde{D}_m$ is a differential satisfying these requirements so that its divisor is $(\tilde{\theta}) = \tilde{P}_i{}^{m_i} \tilde{Q}_j{}^{n_j}$, $1 \leqq i \leqq t$, $n_j \geqq 0$, $m_i \geqq (\tilde{P}_i, m) - m$, and $d(\tilde{\theta}) = \sum_i m_i + \sum_j n_j = 2m(\tilde{g} - 1)$. Then all other differentials $\tilde{\theta}_1 \in \tilde{D}_m$ with this property are of the form $\tilde{f}\tilde{\theta}$ where $\tilde{f}$ is a function on $\tilde{S}$ whose divisor is a multiple of the divisor $b = \tilde{P}_i{}^{(\tilde{P}_i, m) - m - m_i} \tilde{Q}_j{}^{-n_j}$. The space of such functions has dimension $r(b) = d(b^{-1}) + i(b^{-1}) + 1 - \tilde{g}$. Now

$$d(b^{-1}) = \sum_i m_i + \sum_j n_j + \sum_{i=1}^{t} (m - (\tilde{P}_i, m)) = 2m(\tilde{g} - 1)$$
$$+ \sum_{i=1}^{t} (m - (\tilde{P}_i, m)) > 2m(\tilde{g} - 1 > 2g - 2$$

if $g > 1$ so that $i(b^{-1}) = 0$ while if $g = 0$ or 1, $i(b^{-1}) = 0$ because then there are never any differentials in $\tilde{A}_1$ with zeros. Thus we have that

$$\dim A_m{}^H = r(b) = (2m - 1)\,(\tilde{g} - 1) + \sum_{i=1}^{t} (m - (\tilde{P}_i, m)).$$

Now this analysis was predicated on the assumption that there was a non zero $\tilde{\theta} \in \tilde{D}_m$ having the desired properties so that all others could be obtained as $\tilde{f}\tilde{\theta}$. Clearly if $\tilde{g} \geqq 1$ then since $\tilde{A}_m \subset \pi(A_m{}^H)$ and $\dim \tilde{A}_m \geqq 1$ this requirement is met. If $\tilde{g} = 0$ this may not be so. For example, if $\tilde{g} = 0$ then $t \geqq 3$ is an immediate consequence of the Riemann-Hurwitz relation. Also in this case $2m(\tilde{g} - 1) = -2m$ so that any differential of type m on the sphere has at least $2m$ poles. But if $t = 3$, $m = 2$ then any differential in $\pi(A_2{}^H)$ has at most three poles which is impossible, hence $\pi(A_m{}^H) \cong 0$ and $\dim A_m{}^H = 0$.

If $t \geqq 3$, $m \geqq 2$ it will in general depend on the properties of the covering, i.e. the numbers $(\bar{P}_i, m)$, whether $\dim A_m{}^H > 0$ or not. We summarize our results in the following theorem.

THEOREM 3. (a) *If H has no fixed points then*

$$\pi\colon A_m{}^H \cong \tilde{A}_m \text{ and } \dim A_m{}^H \geqq 2m - 1 + \delta_{1m}.$$

Assuming now $\tilde{S}$ has $t > 0$ branch points, then

(b) $\pi\colon A_1{}^H \cong \tilde{A}_1$. *In particular* $\dim A_1{}^H = \tilde{g}$ *so that there are invariant abelian differentials of the first kind different from zero if and only if $\tilde{S}$ is not the sphere.*

(c) *If $m > 1$ and $\tilde{g} \neq 0$, then*

$$1 \leqq \dim A_m{}^H = (2m - 1)(\tilde{g} - 1) + \sum_{i=1}^{t} (m - (\bar{P}_i, m)).$$

If $\tilde{g} = 0$ then $\dim A_m{}^H$ may be zero but in all cases is given by max

$$(0, 1 - 2m + \sum_{i=1}^{t} (m - (\bar{P}_i, m))).$$

The only point needing comment is (a) where $\dim A_m{}^H \geqq 2m - 1 + \delta_{1m}$ is stated. This is so, for, $\tilde{g} \geqq 2$ in the case of a smooth covering. This may be seen from the Riemann-Hurwitz relation or by observing that since $g \geqq 2$ its universal covering surface is the disk and hence if $\pi\colon S \to \tilde{S}$ is smooth $\tilde{S}$ admits the disk as universal covering surface which implies $\tilde{g} \geqq 2$.

3. We now wish to calculate explicitly the matrix representation of R_m, at least for a special case. That is, consider an element $h \in H(S)$ of prime order N and let H be the cyclic group generated by h. Our first observation is that for suitable choice of basis in A_m, R_m has diagonal form. Indeed, since $R_m{}^N = I$, its minimal polynomial $M(\lambda)$ divides $\lambda^N - 1$, hence has distinct roots. A known result of linear algebra is that a matrix can be diagonalized if its minimal polynomial has distinct roots. Another way is to recall two elementary facts of the theory of representations. Namely, every irreducible representation of an abelian group is one dimensional and that every representation of a finite group is completely reducible into a sum of irreducible representations. Thus the matrices of a representation of any abelian group can be brought simultaneously into diagonal form.

The irreducible one dimensional representations of the cyclic group H of order N are simply obtained by $h \to \mu^j$ $0 \leqq j \leqq N - 1$ where $\mu = e^{2\pi i/N}$. In the diagonal form of $R_1(h)$ let us denote by n_j the multiplicity of μ^j. Thus

1 occurs n_0 times and $\sum_{j=0}^{N-1} n_j = g$. But clearly 1 occurs with a multiplicity equal to the dimension of the space $A_1{}^H$ of H invariant differentials which by Theorem 3 gives $n_0 = \tilde{g}$, the genus of S/H. The Riemann-Hurwitz relation informs us immediately that $\tilde{g} < g$ so that for some index k, $1 \leqq k \leqq N-1$, $n_k \neq 0$. (This incidentally is another proof that R_1 is faithful.) We now make use of the simplifying assumption that N is prime by noting that in this case H has no proper subgroup so that any point fixed under an element of H is left fixed by the whole group H. Hence in the covering $\pi\colon S \to \tilde{S} = S/H$ every one of the $t \geqq 0$ branch points $P_1, \cdots, P_t \in S$ are of order $N-1$ and the Riemann-Hurwitz relation is $2g - 2 = N(2\tilde{g} - 2) + (N-1)t$. Now $n_k \neq 0$ means there is a $\theta \in A_1$ such that $h(\theta) = \mu^k\theta$. This implies that if $Q \in S$, not a branch point, is a zero of θ of order q then each of the points $h^j(Q)$, $1 \leqq j \leqq N-1$, is a zero of order q. Thus the divisor of θ has the form

$$(\theta) = P_1{}^{m_1} \cdots P_t{}^{m_t} \cdots Q_j{}^{q_j} h(Q_j)^{q_j} \cdots h^{N-1}(Q_j)^{q_j} \cdots$$

$m_i \geqq 0$, $q_j \geqq 0$, where $\sum_{i=1}^{t} m_i + N \sum_{j} q_j = 2g - 2$. (If $t = 0$, any expressions involving indices running from 1 to t are simply omitted.) Note that q_j may be arbitrary but not the m_i. In fact if z is a local parameter at a branch point P such that h^{-1} is locally $z \to \eta z$, η a primitive N-th root of unity and θ locally $(a_0 + a_1 z + \cdots)\,dz$ then the condition $h(\theta) = \mu^k\theta$ says that

$$(a_0 + a_1(\eta z) + \cdots)\eta\,dz = \mu^k(a_0 + a_1 z + \cdots)\,dz.$$

Now if $\mu^k = \eta^p$, $1 \leqq p \leqq N-1$, then the only non zero coefficients a_n possible are when $a_n\eta^{n+1} = a_n\mu^k = a_n\eta^p$, or $n + 1 \equiv p \pmod N$. In particular the first non zero coefficient has index of the form $rN + p - 1$, $r \geqq 0$ an integer. Now of course at each branch point P_i there is an η_i, p_i, r_i which needn't be the same. Note that the η_i, p_i do not depend on θ while r_i does. Nevertheless we have that the order m_i of the zero of θ at P_i is of the form $m_i = r_iN + p_i - 1$. We are now in a position to apply Riemann-Roch. For if ϕ is another differential satisfying $h(\phi) = \mu^k\phi$ then $\phi/\theta = f$ is an H invariant function on S with poles at most of order q_j at the non branch points Q_j, $h(Q_j), \cdots, h^{N-1}(Q_j)$ while at any branch point $P_i f$ has at most a pole of order r_iN, for ϕ, by what we just noted above, has at P_i a zero of the form $r_i{}^*N + p_i - 1$, $r_i{}^* \geqq 0$. On the other hand any H invariant f with these properties gives $\phi = f\theta$ satisfying $h(\phi) = \mu^k\phi$. Hence n_k is the dimension of the space of such H invariant functions f. By the analysis of the previous section these may be lowered to functions $\tilde{f} = \pi(f)$ on $\tilde{S}$ and these are then multiples of the divisor

$b = \tilde{P}_1^{-r_1} \cdots \tilde{P}_t^{-r_t} \cdots \tilde{Q}_j^{-q_j} \cdots$, and $n_k = r(b) = d(b^{-1}) + i(b^{-1}) + 1 - \tilde{g}$; $d(b^{-1}) = \sum_{i=1}^{t} r_i + \sum^{j} q_j$. We know $\sum_{i=1}^{t} (r_i N + p_i - 1) + N \sum_j q_j = 2g - 2$ so that

$$d(b^{-1}) = \frac{2g-2}{N} - \frac{\sum_{i=1}^{t}(p_i - 1)}{N}.$$ Since

$$1 \leqq p_i \leqq N-1,\ (N-2)t \geqq \sum_{i=1}^{t}(p_i - 1) \geqq 0,$$

and $\frac{2g-2}{N} \geqq d(b^{-1}) \geqq \frac{2g-2}{N} - \frac{(N-2)}{N} t$. There are now two possibilities. If $t = 0$, the $d(b^{-1}) = \frac{2g-2}{N} = 2\tilde{g} - 2$ and $i(b^{-1}) = 0$ or 1. But if $i(b^{-1}) = 1$ there is a differential $\tilde{\Psi} \in \tilde{A}_1$ which upon lifting to $\Psi = \pi^{-1}(\tilde{\Psi}) \in A_1^H$ has zeros precisely at the zeros of θ but $\Psi \neq \theta$ since $h(\Psi) = \Psi$ while $h(\theta) = \mu^k\theta$, $k \neq 0$. Thus $i(b^{-1}) = 0$. If $t > 0$ then

$$d(b^{-1}) \geqq \frac{2g-2}{N} - \frac{(N-2)}{N} t > \frac{2g-2}{N} - \frac{(N-1)}{N}\ t = 2\tilde{g} - 2$$

and so we obtain directly that $i(b^{-1}) = 0$. In any event, we have that $n_k = r(b) = 2\tilde{g} - 2 + 1 - \tilde{g} = \tilde{g} - 1$ in the case that $t = 0$ while if $t > 0$ there is an inequality $\frac{2g-2}{N} + 1 - \tilde{g} \geqq n_k \geqq \frac{2g-2}{N} - \frac{(N-2)}{N}\ t + 1 - \tilde{g}$ which may be written as $\tilde{g} - 1 + \frac{(N-1)}{N} t \geqq n_k \geqq \tilde{g} - 1 + t/N$. Again, if $t = 0$, let us assume that for indices $k_1, \cdots, k_r$, $1 \leqq k_i \leqq N-1$, $1 \leqq i \leqq r$, $n_{k_i} \neq 0$ while otherwise $n_k = 0$. Since $\sum_{j=0}^{N-1} n_j = g$ we have that $\tilde{g} + r(\tilde{g} - 1) = g$. Comparison with the Riemann-Hurwitz relation $2g - 2 = N(2\tilde{g} - 2)$ and recalling that $\tilde{g} \geqq 2$ if $t = 0$ (see remarks following Theorem 3) yields $r = N - 1$. We summarize our results as

THEOREM 4. *If $h \in H(S)$ has prime order N and $t \geqq 0$ fixed points and n_j is the multiplicity of μ^j in the diagonal form of $R_1(h)$, $\mu = e^{2\pi i/N}$, $0 \leqq j \leqq N-1$, then*

(a) *$n_0 = \tilde{g} =$ genus of S/H, H being the cyclic group generated by h.*

(b) *If $t = 0$, $n_j = \tilde{g} - 1$ for $1 \leqq j \leqq N-1$.*

(c) *If $t > 0$, there is at least one index k, $1 \leqq k \leqq N-1$, such that $n_k \neq 0$, and for any such index,*

$$\tilde{g} - 1 + \frac{(N-1)}{N}\ t \geqq n_k \geqq \tilde{g} - 1 + \frac{t}{N}.$$

A similar analysis can be done for the representation matrix $R_m(h)$ for $m > 1$ and the particular case $m = 2$ has been examined elsewhere [5]. In this connection the reader may consult the article of Rauch [7] where certain of these formulas are utilized.

4. Fix $m \geqq 1$ and let $k = \dim A_m = (2m-1)(g-1) + \delta_{1m}$, we are assuming S has genus $g \geqq 2$ in what follows. If $\theta_1, \cdots, \theta_k$ is an arbitrary basis for A_m and z a local parameter at $P \in S$, let $r(t)$, for t an integer $\geqq 1$, denote the rank of the matrix

$$\begin{pmatrix} \theta_1(z), \cdots, \theta_k(z) \\ \theta_1'(z), \cdots, \theta_k'(z) \\ \theta_1^{(t-1)}(z), \cdots, \theta_k^{(t-1)}(z) \end{pmatrix}.$$

One verifies easily that $r(t)$ is independent of both the local parameter z and the choice of basis $\theta_1, \cdots, \theta_k$. In fact $k - r(t)$ is the dimension of the subspace of A_m vanishing with order $\geqq t$ at P. Thus $r(t) = k$ for $t > 2m(g-1)$, for any $\theta \neq 0$ in A_m has only $2m(g-1)$ zeros, while $r(1) = 1$, for not all A_m can vanish at a point. Also, $r(t) = r(t-1) + 1$ if and only if there is a $\theta \in A_m$ with zero of order $t-1$ at P. As t goes from 1 to $2m(g-1)+1$, $r(t)$ changes from 1 to k. If we agree to set $r(0) = 0$, then we see that there are precisely k positive numbers t_j, $1 \leqq j \leqq k$ such that $r(t_j) = r(t_j - 1) + 1$, and these satisfy $1 = t_1 < t_2 < \cdots < t_k \leqq 2m(g-1)+1$, $t_j \geqq j$. Let us call this sequence of k numbers the sequence of m-gaps at P and denote it as $\gamma^m(P) = (\gamma_1^m(P), \cdots, \gamma_k^m(P))$. We have then, by definition, that a number γ is an m-gap at P if and only if there exists a $\theta \in A_m$ with a zero of order $\gamma - 1$ at P. For $m = 1$, one obtains the gap sequence defined in the first section. Here also, as in the case $m = 1$, one may show (see Hurwitz [4]) that there are only a finite number of points for which $\gamma_k^m > k$, and these may be called the m-Weierstrass points.

Using the m-gap sequence one may construct a normalized basis for A_m at P with respect to a given local parameter z as follows. For convenience, since m, P are fixed, let $\gamma_1, \cdots, \gamma_k$ be the m-gap sequence at P and set $\mu_i = \gamma_i - 1$. Let for $1 \leqq i \leqq k$, $\phi_i \in A_m$ have a zero of order μ_i at P. Then in the parameter z,

$$\phi_i = (a_{\mu_i}^{(i)} z^{\mu_i} + \cdots) dz^m, \qquad a_{\mu_i}^{(i)} \neq 0.$$

The differentials ϕ_i are clearly linearly independent and being k in number, form a basis of A_m. Then the differentials $\psi_i = (a_{\mu_i}^{(i)})^{-1}\phi_i - \sum_{j>i} a_{\mu_j}^{(i)} (a_{\mu_j}^{(j)})^{-1} \phi_j$ are again a basis of A_m and satisfy the following normalization:

$$\psi_i = (b_{\mu_i}^{(i)} z^{\mu_i} + \cdots) dz^m,\ b_{\mu_i}^{(i)} = 1, \text{ and } b_{\mu_j}^{(i)} = 0 \text{ for } j > i.$$

The basis $\psi_1, \cdots, \psi_k$ is uniquely determined by these conditions and will be called the basis for A_m at P, z.

If now h is an automorphism of S and $P = h(P)$ is a fixed point, consider $R_m(h)$ with respect to the basis ψ_i at P. Since $R_m(h)(\psi) = \psi h^{-1}$, if h^{-1} is locally $z \to \epsilon z$, one has

$$R_m(h)(\psi_i) = ((\epsilon z)^{\mu_i} + \cdots + b_n^{(i)}\epsilon^n z^n + \cdots)\epsilon^m dz^m.$$

On the other hand, $R_m(h)(\psi_i) = \sum_{j=1}^{k} c_{ij}\psi_j$, where (c_{ij}) is a non-singular matrix. Since $R_m(h)(\psi_i)$ has the same order zero at P as does ψ_i we must have $c_{ij} = 0$ for $j < i$, while by the above expansion for $R_m(h)(\psi_i)$ and the normalization of the ψ_i, we see that $c_{ii} = \epsilon^{\mu_i + m}$, $c_{ij} = 0$ for $j > i$. Thus the matrix $R_m(h)$ is obtained explicitly as a diagonal matrix. We summarize our results as follows.

THEOREM 5. *Let $\gamma_1, \cdots, \gamma_k$ be the m-gap sequence at P and suppose $P = h(P)$ for an automorphism h of S. Furthermore, let h^{-1} at P be given by $z \to \epsilon z$, where ϵ is some primitive N-th root of unity, N being the order of h. Then with respect to the normalized basis of A_m at P, z, $R_m(h)$ is the $k \times k$ diagonal matrix, with diagonal $(\epsilon^{\gamma_1+m-1}, \cdots, \epsilon^{\gamma_k+m-1})$. In particular, the character X_m of the representation R_m is*

$$X_m(h) = \operatorname{trace} R_m(h) = \epsilon^{m-1} \sum_{i=1}^{k} \epsilon^{\gamma_i}.$$

The importance of this theorem is that it gives $R_m(h)$ in terms only of the rotation of h at one fixed point P, and the m-gaps at P. On the other hand, Chevalley and Weil [2] have given a formula for the multiplicity of a given irreducible representation in R_m which depends on the behavior of h at all the fixed points. A combination then of these two formulas enables one to obtain information, sufficient in certain cases, to explicitly calculate gap sequences. In this connection, see [6].

For an application of Theorem 5, take $m = 1$, h an automorphism of prime order N, having $t > 0$ fixed points. Let H be the cyclic group generated by h, and assume there is a fixed point $P = h(P)$ which is not a Weierstrass point, $P \notin W$, so that $\gamma_j = j$, $1 \leqq j \leqq g$. Now if $g = qN + r$, $q \geqq 0$, $0 \leqq r \leqq N$, is the result of dividing g by N then in the diagonal $(\epsilon^1, \cdots, \epsilon^g)$, 1 will occur exactly q times as $\epsilon^N, \epsilon^{2N}, \cdots, \epsilon^{qN}$. But we have already ascertained (Theorem 4) that the multiplicity of 1 in the diagonal is $\tilde{g}$. Thus $g = \tilde{g}N + r$. Comparing this expression for g with the Riemann-Hurwitz relation, $2g - 2$

$=N(2\tilde{g}-2)+(N-1)t$ gives $t=2+2,(\frac{r}{N-1})$. Since t is an integer we must necessarily have that either $r=0$, $r=\frac{N-1}{2}$, $r=N-1$, which give corresponding values for t as $t=2$, $t=3$, $t=4$. This yields

THEOREM 6. *If $P=h(P)$, $P\notin W$, then there are at least 2 and at most 4 fixed points under h.*[2] *The genus $\tilde{g}$ of S/H is given by $\tilde{g}=[g/N]$, the greatest integer in g/N. Writing $g=\tilde{g}N+r$ there are only three possible cases:*

(a) $r=0$, $g=\tilde{g}N$, $t=2$.

(b) $r=\dfrac{N-1}{2}$, $g=(\tilde{g}+\frac{1}{2})N-\frac{1}{2}$, $t=3$.

(c) $r=N-1$, $g=(\tilde{g}+1)N-1$, $t=4$.

Since the theorem assures us of the existence of another fixed point, say P_1, let $(\gamma_1,\cdots,\gamma_g)$ be the gaps at P_1 and suppose h^{-1} at P_1 is locally $z\to\eta z$. Then—understanding equality of diagonal matrices, in the present context, modulo the order of the terms—we must have

$$R_1(h)=\operatorname{diag}(\epsilon^1,\cdots,\epsilon^g)=\operatorname{diag}(\eta^{\gamma_1},\cdots,\eta^{\gamma_g}).$$

This generally should yield some information about the γ_j. However, in the general case where the γ_j may be quite arbitrary it seems that nothing more precise can be said. Also, we wish to consider the case where we are given only that $h(P)=P\in W$, not covered by Theorem 6. For specific results let us consider the following two cases which have the simplification that the gap sequence is known for a point as soon as it is known whether or not that point is in W or not. I. S hyperelliptic: there are $\omega=2g+2$ points in W at each of which the gap sequence is $1,3,5,\cdots,2g-3,2g-1$. II. S is at the opposite extreme, i.e. there are $\omega=(g-1)g(g+1)$ points in W, each having gap sequence $1,2,3,\cdots,g-1,g+1$. From now on we denote by $\operatorname{diag}(a_1,\cdots,a_n)$ an $n\times n$ diagonal matrix with these elements as diagonal.

I. S is hyperelliptic. Let J be the unique involution which leaves W fixed. Since J has order 2, at any point $P\in W$, $J=J^{-1}$ is locally $z\to -z$, $\epsilon=-1$ and so

$$R_1(J)=\operatorname{diag}((-1),(-1)^3,\cdots,(-1)^{2g-1})=\operatorname{diag}(-1,-1,\cdots,-1).$$

From this, $J(\theta)=-\theta$ for every $\theta\in A_1$. This fact is usually proved by

[2] R. Accola has recently pointed out to me that this result can be obtained directly from the Riemann-Hurwitz relation without Theorem 5.

explicitly calculating the differentials which this method obviates. Now let us prove the following: If h is an involution of S and $h(P)=P$, $P\in W$, then $h=J$. For we have locally at P, $z\to -z$ so

$$R_1(h)=\operatorname{diag}((-1),(-1)^3,\cdots,(-1)^{2g-1}=R_1(J),$$

so that since R_1 is faithful $h=J$. Hence given any involution h on S, either h has no fixed points, or has non W fixed points only, or is J. In case of non W fixed points, the three cases of Theorem 10, since $N=2$, reduce to 1) $g=2\bar{g}$, $t=2$; 2) $g=2g+1$; where $\bar{g}=[g/2]$.

Assume now h is an automorphism of prime order N, such that $2<N<2g$ and say $P=h(P)\in W$ and h^{-1} is locally $z\to\epsilon z$. Then $R_1(h)=\operatorname{diag}(\epsilon,\epsilon^3,\cdots,\epsilon^{2g-1})$. A 1 occurs at $\epsilon^N,\epsilon^{3N},\cdots$. Write $2g-1=qN+r$, $0\leqq r<N$. Then if q is even, 1 occurs $q/2$ times, and if q odd, 1 occurs $\dfrac{q+1}{2}$ times. Hence if q even, $\bar{g}=q/2$ and $2g-1=2\bar{g}N+r$ and if q odd $\bar{g}=\dfrac{q+1}{2}$ so $2g-1=(2\bar{g}-1)N+r$.

1) If $2g=(2\bar{g}-1)N+r+1$, comparing with Riemann-Hurwitz gives $t=1+\dfrac{r}{N-1}$, so that a) $r=0$, $t=1$, $(2g-1)=(2\bar{g}-1)N$ or b) $r=N-1$, $t=2$, $2g=2\bar{g}N$ or $g=\bar{g}N$. Now we claim the case a) cannot occur. To see this we recall the representation of h as a permutation of W called $L(h)$. Written in cycle notation, $L(h)$, since of prime order N and operating on $2g+2$ letters, must consist of say L_1 cycles each of N elements and L_2 cycles of single elements. Hence $L_1N+L_2=g+2$. But clearly $L_2=t_W$, t_W being the number of fixed W points so that $L_1N+t_W=2g+2$. However in a) above $2g-1=(2\bar{g}-1)N$ so that $2g+2=(2\bar{g}-1)N+3$, and since $N>2$, either $t_W=3$ or $t_W=0$ and $N=3$, both of which contradict $t=1$. Also, in b) the second fixed point must also be a W point. This may be seen in two ways. One is as before to write $L_1N+t_W=2g+2$. But here $g=\bar{g}N$, so $2g+2=2\bar{g}N+2=L_1N+t_W$ so that $t_W=t=2$. Another way is to recall that $hJ=Jh$ always so that if $h(P)=P$ then $Jh(P)=hJ(P)$ so that $J(P)$ fixed. Hence if the second fixed point P were not a W point, there'd be a third fixed point $J(P)$, contradicting $t=2$.

2) Consider $2g-1=2\bar{g}N+r$; comparing this with Riemann-Hurwitz we have $t=2+\dfrac{r+1}{N-1}$. Since t is an integer and $0\leqq r<N$, then, necessarily, $r=N-2$ and $t=3$, $2g=(2\bar{g}+1)N-1$.

But then $2g+2=(2\bar{g}+1)N+1$ so that $t_W=1$ or $t_W=N+1>3$

which cannot be, hence $t_W = 1$ and the two other fixed points are not W points. Now $\tilde{g} = 0$ implies $N = 2g + 1$ contrary to our assumption; hence $\tilde{g} \geqq 1$ and $N = \dfrac{2g+1}{2\tilde{g}+1} < \dfrac{2g+1}{3} < g$. Summarizing these results we have:

THEOREM 7. *Let S by hyperelliptic. Any involution on S either has no fixed points, or only non W fixed points or is J. If h has prime order N, $2 < N < 2g$, assume there is a W point fixed. Then there are only the two possibilities:*

1) $g = \tilde{g}N$ and there are two W fixed points and these are all the fixed points.

2) $N = \dfrac{2g+1}{2\tilde{g}+1}$, $\tilde{g} \geqq 1$, $t = 3$, the other two fixed points being P, $J(P)$ for $P \notin W$.

Now in both cases, 1) $N = g/\tilde{g} \leqq g$, 2) $N = \dfrac{2g+1}{2\tilde{g}+1} \leqq \dfrac{2g+1}{3} \leqq \frac{2}{3}g + \frac{1}{3} < g$ we see that a necessary condition that a W point be fixed is $N \leqq g$. But if $g < N < 2g$, there being no fixed W point, then $2g + 2 = L_1 N$, so the only possibility then is $N = g + 1$. Hence we have the

COROLLARY 3. *The only possible h of prime order N, $g < N < 2g$, on a hyperelliptic S, is for $N = g + 1$, prime.*

Finally to complete our survey of the hyperelliptic case, we must consider prime order $N > 2g$. But we must have $2g + 2 = L_1 N + t_W$, so the only possible case is $N = 2g + 1$ and there is one fixed W point P. Thus $R_1(h) = \operatorname{diag}(\epsilon, \epsilon^3, \cdots, \epsilon^{2g-1})$ for some N-th root ϵ; hence 1 does not occur at all so $\tilde{g} = 0$. Then by Riemann-Hurwitz we get $2g - 2 = -2(2g+1) + 2gt$ or $t = 3$ and the other two fixed points are some P and $J(P)$.

We can give an illustration where this last case arises. Consider the surface defined by $y^2 = x^6 - x$ which is of genus 2 thus hyperelliptic. Then one see that $x \to \epsilon x$, $y \to \epsilon^{\frac{1}{2}} y$ where $\epsilon = e^{2\pi i/5}$ is an automorphism of order $5 = 2.2 + 1$. The fixed points are clearly the single W point lying over $x = 0$ and the two corresponding points lying over $x = \infty$, so that the theory is verified. To compute $R_1(h)$, at the point over $x = 0$, y is a local parameter, and so h^{-1} is locally $y \to -e^{-2\pi i/10} y$. Now $-e^{-2\pi i/10} = \epsilon^2$. The gaps are $1, 3 = 2g - 1$. Thus $R_1(h) = \operatorname{diag}(\epsilon^2, (\epsilon^2)^3 = \epsilon)$, from which we also see that $\tilde{g} = 0$ as expected since no 1 occurs. To verify this expression for $R_1(h)$, consider the basis for A_1 given by $\dfrac{dx}{y}, \dfrac{xdx}{y}$. Then $\dfrac{dx}{y} \to \dfrac{\epsilon^{-1}dx}{-\epsilon^{-\frac{1}{2}}y} = \epsilon^2 \dfrac{dx}{y}$ (remember $h(\theta) = \theta h^{-1}$) and $\dfrac{xdx}{y} \to \dfrac{\epsilon^{-2}xdx}{-\epsilon^{-\frac{1}{2}}y} = \epsilon \dfrac{xdx}{y}$ so that $R_1(h) = \operatorname{diag}(\epsilon^2, \epsilon)$.

II. We now consider the surface S with $(g-1)g(g+1)W$ points, with gaps $1,2,\cdots,g-1,g+1$ at each one of these points. Let h be of prime order N. If a W point P is fixed, then for some ϵ,

$$R_1(h)=\operatorname{diag}(\epsilon,\epsilon^2,\cdots,\epsilon^{g-1},\epsilon^{g+1}).$$

Write $g=qN+r$. Then we have the possibilities

1) $r=0$, 1 occurs $q-1$ times, $g=(\tilde{g}+1)N$
2) $r=N-1$, 1 occurs $q+1$ times, $g=\tilde{g}N-1$
3) $0<r<N-1$, 1 occurs q times, $g=\tilde{g}N+r$.

Say now a point $P_1\notin W$ is fixed also. Then we know either $r=0$, $r=N-1$ or $r=\dfrac{N-1}{2}$. But then when $r=0$ we know $g=\tilde{g}N$ and $r=N-1$ implies $g=(\tilde{g}+1)N-1$ (by Theorem 6) which are incompatible with 1) and 2) above. Thus necessarily then $r=\dfrac{N-1}{2}$ $(N>2)$ and for some N-th root η, we must have

$$\operatorname{diag}(\epsilon,\epsilon^2,\cdots,\epsilon^{g-1},\epsilon^{g+1})=\operatorname{diag}(\eta,\eta^2,\cdots,\eta^{g-1},\eta^g).$$

Let us note that if $N=5=2.2+1$, $r=2$, this might indeed occur, if say $\epsilon=\eta^2$, then $\epsilon^{r-1}=\epsilon$ and $\epsilon^{r+1}=\epsilon^3=\eta^6=\eta$ so that $(\epsilon,\epsilon^3)=(\eta,\eta^2)$; also if $N=3$, $r=1$, so we may have $(\epsilon^2)=(\eta)$ if $\epsilon^2=\eta$. Now assume $N>5$ so that $r>2$. Clearly $\epsilon=\eta$ is impossible. Let $\eta^a=\epsilon$ where $1<a\leqq r$. Then $(r-1)a\geqq 2r-2>r$ so that there is an integer b, $1<b\leqq r-1$ such that $r<ab<2r+1$. But then ϵ^b occurs in the set $(\epsilon,\epsilon^2,\cdots,\epsilon^{r-1},\epsilon^{r+1})$, and $\epsilon^b=\eta^{ab}$ which does not occur in the set $(\eta,\eta^2,\cdots,\eta^r)$. Thus if $N>5$, either there are only W or non W points fixed. Now if there are no W points fixed, we must have $(g-1)(g)(g+1)=L_1N$ so that N, being prime, divides either $g-1$, g, or $g+1$. Referring now to Theorem 6, assuming there is a non-W point fixed, we see that in cases (a), (c), $N\mid g, N\mid(g-1)$, resp. However, in case (b) $g=\tilde{g}N+\dfrac{N-1}{2}$, so that if $N>3$, N divides none of $g-1$, g, $g+1$ so that this case is impossible. Now assume there is a point $P\in W$ which is fixed. Consider now again the three possible cases we had before (not those of Theorem 6). In 1) $g=(\tilde{g}+1)N$ and on comparison with Riemann-Hurwitz $2g-2=N(2\tilde{g}-2)+(N-1)t$ we have $t=4+\dfrac{2}{N-1}$ so that (1a) $t=6$, $N=2$, or (1b) $t=5$, $N=3$. Since $t>4$ all the fixed points must be W points, but in (1b) we then have $(g-1)(g)(g+1)=3L_1+5$ which is impossible because one of the three

numbers on the left side is divisible by 3. Now in 2) $g=\tilde{g}N-1$ and again we deduce in the usual way $t=2-\dfrac{2}{N-1}$ and since we're assuming $t\neq 0$, we must have $N=3$, $t=1$, but again by $(g-1)g(g+1)=3L_1+1$, this is seen to be impossible. Finally, in 3) $g=\tilde{g}N+r$, $0<r<N-1$ which gives $t=2+\dfrac{2r}{N-1}$ so that necessarily $r=\dfrac{N-1}{2}$ and $t=3$ and if $N>5$ we've seen these are all W points, so that $(g-1)g(g+1)=L_1N+3$.

Summing up then the main results here, we have

THEOREM 8. *On a surface with $(g-1)g(g+1)W$ points each with gap sequence $(1,2,\cdots,g-1,g+1)$ any automorphism of prime order $N>5$ has either only non W or only W points fixed. If only non W then either $N\mid g-1$ or $N\mid g$ and only cases* (a) *and* (c) *of Theorem* 6 *are possible. If only W, then there are exactly three W points fixed and $g=\tilde{g}N+\dfrac{N-1}{2}$ and $(g-1)g(g+1)\equiv 3\ (\operatorname{mod} N)$. If $N\leqq 5$ and there is a W point fixed, either $N=2$ and there are 6 W fixed points, or $N=3$ and there are 3, not necessarily all in W, fixed points.*

As an example of the use of some of the abstract theory developed above, consider a surface S defined by the polynomial $y^5=(x-a_1)^2(x-a_2)^3(x-a_3)^4$. Both x,y are functions on S. Let $P_0,P_1,P_2,P_3\in S$ be those points $x(P_i)=a_i$ $i=1,2,3$, $x(P_0)=\infty$; these are the only branch points of x each of order 4 and by the Riemann-Hurwitz formula S has a genus $2g-2=-10+16$ or $g=4$. In particular x has only 5 poles at P_0 and y has only 9 poles at P_0 as is easily verified. Also y has precisely 9 zeros, 2 at a_1, 3 at a_2, 4 at a_3.

S admits an automorphism of order 5 by defining for $P=(x,y)\in S$, $h(P)=(x,\epsilon y)$ where ϵ is some 5th root of unity, $\epsilon\neq 1$. It is immediate that P_0, P_1, P_2, P_3, are the only fixed points of h, each of order 4 and that $\pi\colon S\to\bar{S}$ in this case is identical with $x\colon S\to$ sphere. In terms of a local parameter at any P_i, h is locally $z\to\epsilon z$ and $h^{-1}\colon z\to\eta z$ where $\eta=\epsilon^{-1}=\epsilon^4$.

Let $\gamma(P_0)=(\gamma_1,\gamma_2,\gamma_3,\gamma_4)$. Then we know that

$$R_1(h)=\operatorname{diag}(\eta^{\gamma_1},\eta^{\gamma_2},\eta^{\gamma_3},\eta^{\gamma_4}).$$

Let us determine $\gamma(P_0)$. Clearly $f_1=x$ is a function at P_0 with 5 poles. By inspection we obtain

$$f_3=\frac{(x-a_1)(x-a_2)^2(x-a_3)^2}{y^2},\ 7\text{ poles at }P_0$$

$$f_2=\frac{(x-a_1)(x-a_2)(x-a_3)}{y},\ 6\text{ poles at }P_0$$

$$f_4 = \frac{(x-a_1)^2((x-a_2)^2(x-a_3)^3}{y^3}, \text{ 8 poles at } P_0$$

Thus $P_0 \notin W$ and $\gamma(P_0) = (1,2,3,4)$ so that $R_1(h) = \mathrm{diag}(\eta, \eta^2, \eta^3, \eta^4)$. Observe that $\eta^5 = 1$ does not appear, as was to be expected, for $\tilde{S}$ here is the sphere and $\tilde{g} = 0$. Consider now $\gamma(P_1) = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$. We know $\gamma_1 = 1$, $\gamma_4 < 8$ and $R_1(h) = \mathrm{diag}(\eta, \eta^{\gamma_2}, \eta^{\gamma_3}, \eta^{\gamma_4})$. If $\gamma_2 = 3$ the surface is hyperelliptic and so 5 is then a gap, so $\eta^5 = 1$ appears which is impossible, thus $\gamma_2 = 2$ and this forces $\gamma_3 = 3$, $\gamma_4 = 4$. Hence, this clearly holds at P_2, P_3, also, and have the same gap sequence.

One may also proceed in a different way, i. e. determine explicitly $R_1(h)$ by computing the differentials and then find the gap sequence by means of this at P_0, P_1, P_2, P_3. In fact, by inspection, one sees that

$$\theta_1 = \frac{(x-a_1)(x-a_2)^2(x-a_3)^3}{y^4}\,dx,\ \theta_2 = \frac{(x-a_1)(x-a_2)(x-a_3)^2}{y^3}dx$$

$$\theta_3 = \frac{(x-a_2)(x-a_3)dx}{y^2},\ \theta_4 = \frac{dx}{y}.$$

Then $\theta_1 h^{-1} = \eta\theta_1$, $\theta_2 h^{-1} = \eta^2\theta_2$, $\theta_3 h^{-1} = \eta^3\theta_3$ and $\theta_4 h^{-1} = \eta^4\theta_4$ so that $R_1(h) = \mathrm{diag}(\eta, \eta^2, \eta^3, \eta^4)$. From this one now may determine the gaps at each of the fixed points.

Consider the general case $y^N = \prod_{i=1}^{L}(x-a_i)^{m_i}$ where N is prime and we assume $1 \leqq m_i \leqq N-1$ and $M = \Sigma m_i$ is such that $(M,N) = 1$. This is a surface S of genus g where $2g-2 = -2N + (N-1)(L+1)$ or

$$g = \frac{(N-1)(L-1)}{2}.$$

This last relation is by observing that x is a function with branch points of order $N-1$ at the points $P_0, P_1, \cdots, P_L$ where $x(P_i) = a_i$, $i = 1, \cdots, L$ and $x(P_0) = \infty$. y is a function with M poles at P_0 and no others.

It is interesting to observe that the genus depends only on N and L and that the exponent m_i of $(x-a_i)^{m_i}$ is irrelevant. This S admits the automorphism of order N defined by $(x,y) \to (x, \epsilon y)$ where $\epsilon^N = 1$, $\epsilon \neq 1$. Again the fixed points are precisely $P_0, P_1, \cdots, P_L$ and if the gap sequence at P_0 say, can be determined to give $R_1(h)$ we may then as before try to use this to compute the gaps at $P_1, \cdots, P_L$. However, it is not clear that in general this method always gives complete results; for example if N is small with respect to L and g, since the powers of elements in $R_1(h)$ are always reduced $\mathrm{mod}\, N$, a decisive determination of the gap sequence may not be available.

In certain cases, as in the one worked out above, this method is though quite useful.

Finally we can observe that in the case of a surface S determined by such a polynomial, the fixed points $P_0, P_1, \cdots, P_L$ are either all W points or all not W points. For $\frac{1}{x-a_i}$ has only N poles at P_i and x has only N poles at P_0. Hence N is a non gap at each. Now $(N-1)(L-1)=2g$ so $N=\frac{2g}{L-1}+1$ (if $L=1$, $g=0$ which is of no interest here); thus if $L>2$, $N \leqq g+1$ and $N=g+1$ only in the case $L=3$, as considered above. Hence if $L>3$, $N<g+1$ and so each P_i is a W point. Note that this result follows also independently from Theorem 6, for since when $L>3$ there are at least five fixed points none of them can be a non W point for in that case by Theorem 6 there are at most four fixed points. If $L=2$, $N=2g+1$ and this is the only case in which the automorphism of the type under consideration can have as fixed points both W and non W points. For example $y^5=(x-a_1)(x-a_2)$. Here S is of genus 2, hyperelliptic, and if $x(P_0)=\infty$, x has 5 poles at P_0 and y has only 2 poles at P_0. Hence P_0 is a W point. On the other hand $\frac{(x-a_2)}{y^3}$ is a function with only three poles at P_1, $x(P_1)=a_1$, and $\frac{(x-a_1)}{y^3}$ has only 3 poles at P_2, $x(P_2)=a_2$. Thus P_1, P_2 are not W points. In fact $P_2=J(P_1)$ by our previous theory.

REFERENCES.

[1] L. Ahlfors and L. Sario, ***Riemann Surfaces***, Princeton, 1960.

[2] C. Chevalley and A. Weil, "Uber das Verhalten der Integrale 1. Gattung bei Automorphismen des Funktionenkorpers," ***Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg***, vol. 10 (1934), pp. 358-361.

[3] A. Hurwitz, "Uber diejenigen algebraischen Gebilde, welche eindeutige Transformationen in such zulassen," ***Mathematische Werke***, Bd. I, Birkhauser, Basel, 1932, pp. 241-259.

[4] ———, "Uber algebraische Gebilde mit eindeutigen Transformationen in sich," ***ibid.***, pp. 392-430.

[5] J. Lewittes, "Invariant quadratic differentials," ***Bulletin of the American Mathematical Society***, vol. 68 (1962), pp. 320-322.

[6] ———, "Gaps at Weierstrass points for the modular group," ***ibid.***, vol. 69 (1963), pp. 578-582.

[7] H. E. Rauch, "Singularities of the modulus space," ***ibid.***, vol. 68 (1962), pp. 390-394.

[8] G. Springer, ***Introduction to Riemann Surfaces***, Addison-Wesley, 1957.

ERRATA.

Eckford Cohen, *Arithmetical notes,* V. Vol. 83 (1961), pp. 683-697.

The term $O(x^{1/k}\beta(r))$ in (2.4) is unnecessary and should be deleted. (See remark 2.1.) Thus (3.9) and the computation preceding it on page 697 should be omitted.

Morris Newman, *Normal congruence subgroups of the modular group.* Vol. 85 (1963), p. 419. Received date should be given as March 13, 1963.

Arthur J. Schwartz, *A generalization of a Poincaré-Bendixson theorem to closed two-dimensional manifolds.* Vol. 85 (1963), p. 453.

line 3—$T \times M — M$ should be $T \times M \rightarrow M$

line 6—$(T\Omega — \Omega)$ should be $(T\Omega = \Omega)$

line 9—beginning of the line should read fixed point rather than $_1$xed point.